NOVA IMS Information Management School

MDSAA

Master Degree Program in

Data Science and Advanced Analytics

Analyzing the Radio Broadcasting Market in Slovenia

A Power BI Analysis

Nina Urbancic

Project Work

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Analyzing the Radio Broadcasting Market in Slovenia

A Power BI Analysis

by

Nina Urbancic

Project Work presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics.

Supervised by

Mijail Juanovich Naranjo Zolotov, Ph.D., NOVA Information Management School (NOVA IMS)

February, 2024

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisbon, 12.2.2024

ACKNOWLEDGEMENTS

I would like to extend my appreciation to my supervisor, Mijail Juanovich Naranjo Zolotov, whose insightful suggestions, and motivating support were invaluable during the writing of this thesis.

Additionally, I am grateful to our client, who prefers to stay unnamed, for allowing the use of their data, which was essential for this project. Their support and resources were greatly appreciated.

My family and friends deserve special acknowledgment for their understanding, patience, and encouragement during this process.

Lastly, but certainly not least, I am deeply thankful to my parents for their endless support and encouragement throughout my academic endeavors.

ABSTRACT

This thesis focuses on the application of Power BI to analyze the Slovenian radio broadcasting market, focusing on demographic preferences for radio stations. It explores how data analytics can help media outlets in customizing marketing strategies to specific audience segments, enhancing listener engagement and retention. The research uses Power BI for data transformation, visualization, and analysis, presenting a methodical approach to data handling from collection through to visualization. The findings strive to offer media houses strategic insights for competitive advantage in a market increasingly influenced by direct competition as well as digital streaming services. The study contributes to the academic and practical understanding of business intelligence in media analytics, focusing on the importance of data-driven decision-making in maintaining relevance and competitiveness in the digital age.

KEYWORDS

Business Intelligence; Power BI; Radio Broadcasting; Data-Driven; Audience Demographic Analysis; Media Consumption Behavior

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

DAX	Data Analysis Expressions
BI	Business Intelligence

1. INTRODUCTION

1.1. MOTIVATION

Over recent years, more and more companies are realizing the importance of business intelligence and striving towards being data-driven. (Sabharwal & Miah, 2021) Our client is no different. Working with many Slovenian media outlets, the company identified an opportunity to create a detailed report focusing on understanding the behavior and preferences of radio station audiences. In times when radio broadcasting stations face increasing competition from new streaming services, which have significantly changed the media consumption market, this competition highlights the necessity for traditional broadcasters to leverage data analytics to remain competitive and relevant. (Hirschmeier et al., 2019)

With a large variety of radio stations broadcasting to a diverse audience, it is becoming more and more important to identify which stations are most listened to by specific demographic groups. This knowledge enables radio stations, or the media houses they belong to, to customize their marketing strategies more effectively. With a better understanding of which demographics are listening to which stations, their marketing campaigns can be more customized and have a higher impact. In turn, this improves the return on investment of marketing campaigns for radio broadcasting stations. (Marketing Charts, 2016)

Moreover, along with marketing campaigns, radio stations or media houses can also customize their content to better suit their audience, generating higher engagement which leads to a higher retention rate of existing listeners and attracting new ones. Many decisions can be made based on this data, such as choosing which radio shows to host, who the most suitable host would be, when the shows should air, and more. This knowledge can help media houses gain considerable competitive advantage. (Marketing Charts, 2016)

Furthermore, understanding how audience preferences evolve over time is also important for media outlets to better anticipate shifts and changes in audience behavior. With these insights, it is easier to retain a loyal audience. (Marketing Charts, 2016)

In the age of rapidly growing data, making decisions based on accurate analytics is increasingly important. (Sabharwal & Miah, 2021)

1.2. OBJECTIVES

The thesis is based on a project from my workplace involving a Slovenian company that specializes in public opinion surveys and market research. While the company prefers to stay unnamed, it has allowed the use of the project for my master's thesis. The focus of this thesis is to use Power BI technology to analyze and display data, making it practical and insightful.

The enhanced objectives of the project include:

- **To Perform a Comprehensive Demographic Analysis:** The core objective is to carefully analyze demographic preferences for radio stations across Slovenia. This involves analyzing the radio listening habits of various population segments—categorized by gender, age, regional location, monthly income, and educational background—to reveal trends and patterns. This analysis aims to provide media outlets actionable insights, helping them to adapt and evolve their strategies to meet the specific needs and preferences of diverse demographic groups.
- **To Utilize Power BI for Data Transformation and Visualization:** Another primary goal is to take advantage of Power BI's capabilities for the transformation, visualization, and analysis of collected data. This includes using Power Query for data cleaning and transformation, as well as using Power BI's data modeling and visualization tools to create insightful reports. The objective extends to demonstrating how Power BI can convert raw data into actionable insights that are of value to businesses.
- **To Conduct a Methodical Approach to Data Handling:** Starting with data collection—carried out by the client through telephone and online surveys—this project loads, cleans, transforms, and models the data within Power BI. This objective highlights the importance of a structured methodology in handling data from its point of collection to the final stages of analysis and reporting.
- **To Offer Strategic Insights to Radio Broadcasters**: The insights gained from the Power BI analysis are intended to be marketed to Slovenian radio stations by our client. This thesis hopes to showcase how these insights can assist broadcasters in making informed strategic decisions to enhance their market position and relevance in the digital era. The focus is on helping radio stations to adapt to changing audience preferences and therefore improving listener engagement and retention.
- **To Bridge the Research Gap in Media Sector Data Analytics:** This work hopes to fill a gap in existing research and practices regarding the application of data analytics within Slovenia's media sector. It aims to highlight the potential of business intelligence tools like Power BI in the media analytics, showcasing their effectiveness in real-world business scenarios.
- **To Serve as a Guide for Industry Professionals**: The thesis hopes to act as a practical guide for industry professionals interested in leveraging data analytics for strategic advantages. It explores the application of data analytics in understanding and catering to audience needs, enabling more targeted and effective marketing and content strategies.

The methodology of this project follows a structured multi-stage process, beginning with the important phase of data collection. This initial step was fully managed by the client, using both telephone and online surveys to gather data from respondents. This approach ensured a wide and diverse dataset, which reflected the target demographic's radio listening habits and preferences.

Once collected, the data was systematically organized and stored in Excel files, which were then uploaded to a SharePoint folder.

The next stage involved the process of data cleaning and transformation, with the use of Power Query within Power BI. This step was important, as it involved cleaning the raw data—removing inconsistencies, errors, or irrelevant information—to ensure its readiness for in-depth analysis. Power Query's advanced data processing features allowed for a transformation of the collected data into a format suitable for analytical exploration and insight generation.

This approach to data management—from collection through cleaning and transformation—provided a basis for the analysis and visualization tasks that followed. This project allowed us to use a wide spectrum of Power BI's analytical tools to extract meaningful insights and trends from the dataset.

An important part of our methodology was the demographic segmentation of the participants in the survey. This process involved dividing the data into distinct groups based on several key characteristics: age, gender, geographical location within Slovenia, monthly income, and educational background. This step helped us prepare a more detailed analysis, allowing us to customize the insights to specific audience segments and identify unique trends and preferences within each demographic group.

Following the segmentation, the last step of our methodology was the interpretation and analysis phase, where the collected data was examined using Power BI. This final step was dedicated to breaking down the data into actionable insights and clear reports. The power of Power BI was instrumental in this process, offering advanced data visualization tools that helped in explaining patterns and trends within the data.

The goal of this project was to get a deeper understanding of the radio listenership in Slovenia, using Power BI as the business intelligence tool. By breaking down the audience into demographic segments, we aimed to discover specific listening preferences and behaviors that could help radio stations prepare more targeted strategies. The insights gathered from our analysis were intended to be packaged and offered to Slovenian radio stations by our client. This approach not only strived to enrich the stations' understanding of their audience but also to provide them with the data-driven strategies necessary to enhance their programming and marketing efforts.

2. LITERATURE REVIEW

2.1. BUSINESS INTELLIGENCE CONCEPTS

Business Intelligence

Business Intelligence (BI) is defined as a technology-driven approach to analyze data and present actionable information to help executives, managers, and other corporate end users make informed business decisions. Consisting of various phases such as data gathering, data storage, knowledge management, and data analysis, it contributes to the business decision-making process. It does so by utilizing tools, applications, infrastructures, and best practices that provide access to and analysis of information to improve and optimize decisions and performance. (Larson & Chang, 2016)

Business Analytics

Business analytics uses a combination of historical and current data to support strategic decisions within organizations. This is often achieved through a combination of statistical analysis and business expertise. Business analytics can be categorized into dashboards, reports, and ad hoc analysis. Dashboards display key indicators and insights that support decision-making, while reports are often created to answer specific business questions and are usually more detailed and therefore actionable. Ad hoc analysis can span from answering specific queries do creating detailed reports. (Reis & Housley, 2022)

Data Warehouse

A data warehouse serves as a centralized repository for reporting and analysis of data that is formatted and structured specifically for analytical use cases. As an important component of business intelligence, data warehouses support the storage, consolidation, and retrieval of large volumes of data from multiple sources. Data warehouses are designed to sustain complex queries and analysis, enabling organizations to make informed decisions based on historical data. They set up a stable space, separate from daily business databases, to easily collect and examine data from different business areas. (Reis & Housley, 2022)

Data Modeling

Data Modeling is the process of structuring and organizing data and is a crucial aspect of data handling. The data is organized in a way that it can reflect real-world scenarios and answer relevant business questions. A strong data model shows how different data relates to each other and how they interact with each other and within an organization. It promotes useful business workflows and decision-making. Good data models are very important in translating complicated data into meaningful business outcomes, ensuring consistency across various departments. Without a good business model, it is unlikely that actionable insights, such as customer behavior patterns, can be generated. (Reis & Housley, 2022)

Data Visualization in Power BI

Data visualization in Power BI is a process of creating a visualization in a fast and user-friendly way. It involves a drag-and-drop technique, with data fields easily transferable between axis labels, legends, and values. Besides the standard graphs and charts, Power BI's visualization includes interactive visuals and more advanced statistical analysis through "R visuals". The key benefit of using Power BI is the simplicity of use and the interactivity of the insights generated. This allows for fast changes in visualization and a dynamic exploration of data. (Becker & Gould, 2019)

ETL (Extract, Transform, Load)

ETL (Extract, Transform, Load) is a data integration process, whose purpose is to combine information from different data sources into a single, consistent database. This database is then transferred to a data warehouse or another similar system. The beginning of this process dates back to 1970s when databases gained popularity. ETL was created as a way of integrating and preparing data for analysis. Soon, it became the main approach for managing data in data warehousing projects. (IBM, 2024)

Today, ETL serves as a basis for varied data-related tasks, including data analytics and machine learning. This is achieved by using pre-determined business rules to clean and organize data, customized to meet specific business intelligence requirements. These requirements can range from standard reporting to advanced analytics. ETL is most often used for three main purposes: extracting data from systems, improving data quality through data cleaning and standardizing, and loading data into a database or system. (IBM, 2024)

Model Schema: Star Schema and Galaxy Schema

A star schema is defined as a organization model with a central fact table which is surrounded by dimensional tables that are simplified for faster querying. We know two main types of star schemas: simple and complex. A simple star schema includes one fact table and its related dimension tables, whereas a complex star schema consists of multiple fact tables that are connected through their corresponding dimensions. (Sumalatha, 2017)

A galaxy schema is a type of a complex star schema which consists of multiple fact tables that share the same dimensions. It is a combination of different data warehouses, offering a broader data analysis and allowing for more flexibility and scalability across various business areas or processes. (Sumalatha, 2017)

2.2. RADIO BROADCASTING BUSINESS

Radio broadcasting

Radio broadcasting involves the transmission of messages to broad audiences using radio waves. Normally, radio is used to deliver audio content, such as talk shows or music programs. Radio stations transmit content at varying frequencies, also known as wavelengths, as there

is a mathematical relationship between the frequency and wavelength of a radio wave. To receive a radio broadcast, you need a radio set to the corresponding frequency. (Techwalla, n.d.)

Beginnings and Brief History

The history of radio broadcasting traces back to Reginald Fessenden's inaugural radio program in the U.S. in 1906. Following World War I, radio experienced rapid commercialization, with KDKA in Pittsburgh, established in 1920, being regarded as the first commercial radio station. This period saw the growth of radio stations and the rise of advertising as a primary source of revenue. (Britannica, 2024)

Network broadcasting became common as stations connected to form networks, assisting the distribution of programs to wider audiences. Regulatory measures, such as the establishment of the Federal Communications Commission (FCC) in the U.S., were introduced to oversee licensing and prevent monopolies. (Britannica, 2024)

The United Kingdom developed its broadcasting system, ending in the formation of the British Broadcasting Corporation (BBC) as a public service broadcaster. Internationally, efforts were made to coordinate broadcasting activities and address challenges related to frequency allocation through the formation of international organizations. (Britannica, 2024)

In summary, radio broadcasting evolved from its experimental origins to become a significant channel for communication and entertainment, shaping global broadcasting standards and regulations along the way. (Britannica, 2024)

Types of Radio Broadcasting

In the context of radio broadcasting, several types are recognized, including AM, FM, Pirate Radio, Terrestrial Digital Radio, Satellite Radio, and Internet Radio. Each type offers unique features from signal adjustment techniques and broadcast ranges to the legalities and modes of reception. (Agile Broadcast, 2015)

- **AM (Amplitude Modulation):** The original radio broadcasting technology. It adjusts a signal's amplitude to transmit over medium-wave and, in some regions like Europe, long-wave bands. Known for its straightforwardness and the ability to produce clearly detectable sound, AM broadcasting has been a reliable means of global radio communication. (Agile Broadcast, 2015)
- **FM (Frequency Modulation):** Uses VHF waves to offer enhanced sound clarity and stereo broadcasting capabilities. This type of broadcasting is common due to its superior audio quality, although it has a more limited range compared to AM, making it ideal for local stations. FM radio's rise in accessibility post-World War II decreased AM radio's dominance, as FM frequencies became more widely available to the public. (Agile Broadcast, 2015)

- **Pirate Radio:** Typically involves unauthorized or unlicensed broadcasts, often for political, entertainment, or communication purposes. Its history includes use by maritime vessels as a broadcasting platform. The legality of these transmissions varies globally, encompassing AM, FM, and shortwave signals. Interestingly, a broadcast's legality can differ between the origin and reception points, especially when crossing international borders. (Agile Broadcast, 2015)
- **Terrestrial Digital Radio:** Also known as DAB Digital Radio (Digital Audio Broadcasting), emerged in the mid-1990s in Europe. It is an advanced system designed to transmit high-quality digital audio and data services to mobile, portable, and fixed receivers via terrestrial transmitters in the VHF bands. It optimizes spectrum and power efficiency in network planning through Single Frequency Network (SFN) and gap-filling strategies, ensuring compatibility with existing radiocommunication services. As a leading standard for digital radio broadcasting, DAB enhances broadcast quality, meets listeners' expectations for sound quality, and reduces energy usage for broadcasters. (WorldDAB, 2024)
- **Satellite Radio**: Uses satellites for broadcasting, enabling coverage over broad geographical areas. Primarily designed for vehicular use, satellite radios offer subscription-based services, ensuring revenue generation. Ground stations transmit signals to satellites, which then transmit broadcasts to receivers in vehicles and homes. These signals, featuring encrypted broadcasts and metadata, require specialized receivers for sorting and display. This technology allows nationwide and global radio broadcasting capabilities. (Agile Broadcast, 2015)
- **Internet Radio:** Marks a significant development in broadcasting, enabled by new communication technologies. It operates in two distinct formats: one where terrestrial radio broadcasts are extended to the Internet, allowing simultaneous transmission over traditional airwaves and digital platforms. This dual-channel approach is particularly appealing to advertisers, as it broadens the reach of broadcasts to include areas beyond the conventional geographic limits of terrestrial signals, therefore offering global accessibility. The second format consists of radio stations that exclusively broadcast online through digital channels. These stations, while lacking the advantages of terrestrial broadcast infrastructure, are attractive to advertisers due to their capacity to provide more accurate and globally accessible publishing and measurement outcomes. This digital-exclusive approach signifies a transformative shift in radio broadcasting, leveraging the Internet to extend its reach and impact beyond traditional boundaries. (Krishnan & Chang, 2000)

Each type signifies a unique approach to signal transmission and audience reach, reflecting radio's adaptability over time. (Agile Broadcast, 2015)

Radio advertising

Radio advertising is a way to promote products, services, or brands using radio waves. It involves broadcasting messages to potential customers through radio commercials, which can be short audio ads featuring jingles, voiceovers, or sound effects. (Singh, 2023)

To reach the right audience, businesses select radio stations that match their target demographics. For example, if a company wants to reach young adults, they might advertise on a station that plays pop music. In summary, radio advertising is an effective way to connect with a large audience and promote products, services, or brands. (Singh, 2023)

2.3. INTEGRATED ANALYSIS OF RELATED WORKS

In this section, I review a series of studies that deepen our understanding of digital innovation in the radio broadcasting and advertising sectors. This study looks at the details of each research and shows how they all connect to the main goal: using digital analytics and technology to improve how we engage audiences and deliver content more efficiently in radio broadcasting.

- **Digital Analytics in Advertising and Broadcasting:** The study by Mehanović & Durmić, 2022 explores the application of business intelligence in digital advertising, highlighting the strategic use of data analytics to optimize ad targeting and enhance user engagement. This work matches the objectives of my thesis by showing the power of analytics in personalizing content, but within the context of digital advertising. The use of the tribeOS platform shows how analytics can save resources and make campaigns more effective. Similarly, my study uses data analytics, especially with Power BI, to improve marketing strategies for radio broadcasters. This aims to better understand what the audience likes and how they behave.
- **Challenges and Adaptations in Digital Broadcasting:** Hirschmeier et al., 2019 explore the digital transformation challenges faced by the radio broadcasting industry, emphasizing the need for adaptation to new technological markets and consumer expectations. This research supports my thesis by showing the importance of using digital tools and strategies to stay competitive and relevant. Looking at expert opinions through qualitative content analysis adds to the discussion on adapting to digital, sharing ideas on how to overcome challenges and innovate in the radio industry. My thesis takes this further by demonstrating how Power BI can be a game-changer in analyzing listener data to customize content and strategies more effectively.
- **Interactive Radio Content and Multimedia Analytics:** Bailer et al., 2020 examine how the changing radio broadcasting market works with multimedia analytics to create interactive content that engages listeners more than just audio. This study is relevant to my work because it shows how data analytics and multimedia tools can improve the radio experience. The focus on making content interactive and customized to what

listeners like and their situations reflects my thesis's goal to use Power BI for increasing listener involvement and keeping them engaged through strategies based on data.

- **Digital Convergence and Transformation of Radio:** The study by Kuyucu, 2019 on digital convergence within the radio sector adds a critical dimension to this integrated analysis. This points out how traditional radio has moved into the digital media world, focusing on its historical development and the digital strategies that have transformed radio broadcasting. This study emphasizes the need to switch to digital platforms for delivering content and engaging audiences, which aligns with my thesis's emphasis on applying data analytics and Power BI to make strategic enhancements in radio broadcasting.
- **Digital Broadcasting – Challenges and Opportunities for European Community Radio Broadcasters:** Hallett & Hintz, 2010 discuss the impact of the digital transition on community radio broadcasters in Europe, highlighting the challenges posed by digital switch-over policies and the opportunities digital technologies may provide for community radio. This study highlights the need for policy adaptations that support the position of community radio in the digital era, reflecting on the broader implications of digitalization for media diversity and accessibility.
- **Transformation of Radio Technology in The Digital Age:** The study by Ismandianto et al., 2022 explores the transformation of radio in response to digital media advancements and the COVID-19 pandemic, focusing on Radio Republik Indonesia's strategies for maintaining relevance. It emphasizes the need for radio to innovate and update its technology to succeed in the digital era, focusing on how integrating digital approaches influences listener participation and the delivery of content. This study highlights the importance of digital platforms in expanding radio's engagement and interactions with its audience, aligning with the thesis's focus on leveraging data analytics and digital strategies in radio broadcasting.

3. METHODOLOGY

The Methodology section of my thesis explains the strategies implemented to achieve the objectives outlined in the Objectives section, detailing the methods deployed to attain these goals and produce outcomes.

3.1. TOOLS AND CONCEPTS

SharePoint

SharePoint is a collaborative platform from Microsoft that's used to create websites. It provides a secure space for storing, organizing, sharing, and accessing information from any device with web browser access, such as Microsoft Edge, Internet Explorer, Chrome, or Firefox. SharePoint is versatile and can be used as part of Microsoft 365 or as a standalone service, allowing for the creation of sites to share documents and information internally within an organization or externally with partners and customers.

Power BI

Power BI is a tool developed by Microsoft that combines various software services and applications. Its primary function is to turn different kinds of data, whether from simple Excel files or complex data warehouses, into clear and visually appealing insights. It allows users to easily connect to these data sources, helps in visualizing key information, and provides options to share these insights with others, supporting data-driven decision-making and collaboration.

Power Query

Power Query is a tool designed for transforming and preparing data. It comes with a user-friendly graphical interface to collect data from a wide range of sources, and a Power Query Editor for making data transformations. Power Query is capable of carrying out ETL (extract, transform, load) tasks and is incorporated into several Microsoft products, including Power BI and Excel. This tool makes the data preparation process much easier, which is a crucial aspect of business analytics, and enables consistent and repeatable operations for modifying and analyzing data.

DAX

DAX, short for Data Analysis Expressions, is a formula language for creating custom calculations in Analysis Services, Power BI, and Power Pivot in Excel. It includes functions, operators, and values for performing calculations and queries on data within tables and columns in tabular models. DAX is integral for measures, calculated columns, tables, and for applying row-level security in data models.

M Code

M Code, also known as Power Query Formula Language, is a data manipulation language used in Microsoft Power BI, Excel, and other Microsoft Power Platform products for data transformation and preparation tasks. M is a functional, case-sensitive language that allows users to import, filter, clean, aggregate, and perform other data transformation operations step-by-step.

Field Parameters

Field parameters in Power BI are a feature that enhances the dynamic capabilities of reports, allowing report users to dynamically control which fields are used in visuals without the need for complex DAX formulas or multiple pages/slicers for different scenarios. This feature significantly improves the flexibility and interactivity of Power BI reports.

Demographics

Demographics are the various characteristics that define different groups within a population, categorized by criteria such as age, gender, income, and more. This categorization helps in analyzing the specific traits of a group. The study of demographic information is important for businesses, organizations, and governments because it supports making well-informed choices. Companies use this information to create targeted marketing plans, and governments use it to design and apply public policies.

Numerus

The term refers to the number of times something happens, appears, or is counted in a dataset. It's often used to indicate how many people or things are studied in research, showing the total amount of participants or observations looked at.

Population

The population refers to the complete set of all possible individuals, observations, or data points that meet a specific set of criteria. The population is the total group that researchers are interested in learning about or drawing conclusions from.

Sample

A sample is a subset of the population that is selected for the actual study or analysis. The sample represents a portion of the population, chosen through various methods (random sampling, stratified sampling, etc.) to reflect the characteristics of the larger group as accurately as possible.

3.2. DATA COLLECTION

The data collection for this study was entirely managed by the client, who specializes in public opinion surveys and market research. This involved their team conducting detailed surveys of

demographic preferences for radio stations in Slovenia through online and telephone surveys. The company's CEO, who preferred to remain anonymous, highlighted their approach to research as *“a carefully orchestrated operation.”*

Their strategy for collecting data guarantees that they thoroughly cover and accurately represent the target population. Surveys are carried out through a balanced mix of phone and online methods to achieve a significant sample size of exactly 288 surveys a day, which amounts to 2,016 surveys a week. The surveys are carried out once every three months, except for May 2023 which was specifically requested and sponsored by an individual media house. This schedule is designed to capture a broad spectrum of listening habits across different days, enhancing the reliability of the data. This ensures that the data is representative for each day of the week.

Phone sample (50%): Uses a two-step random selection, ensuring regional representation based on a universal telephone directory. Households are selected across Slovenian statistical regions, with individuals aged between 15 and 80 years chosen randomly within those households. The sample includes landline and mobile numbers, with mobiles comprising about 30%.

Online sample (50%): Includes members from the company's online panel, based on their internal database, and is enhanced by an additional random online sample, which serves as an addition to the sample on the online panel. The online panel represents 35% or more of the sample and a random online sample up to 15%.

Extra care is taken to focus on the sample's demographic structure, including gender, age, education, and geographic area, to closely mirror the overall population's distribution. To make sure the sample accurately reflects the broader population, a method of adjusting the data through weighting is used.

The survey's goal is to find out which radio channels participants listened to the previous day and the specific times they tuned in. Weekends are included in the survey.

The telephone survey uses a standardized format, beginning with open-ended questions about listening habits from the previous day, followed by socio-demographic questions. The online survey follows a similar structure but replaces open questions with a rotating selection of radio stations for respondents to choose from. Additionally, it includes verification questions and implements security measures to ensure the accuracy of responses and prevent the misuse of the survey, maintaining the integrity of the data collection process.

The client provided me with the telephone version of the questionnaire, which is included in Appendix A of this document.

The data is gathered using the client's internal tool, which provides an in-depth analysis of demographics and geography throughout Slovenia. This analysis focuses on radio audience

sizes, breaking down metrics like reach percentages, actual numbers of listeners (both daily and weekly), ratings in percentages and actual numbers, average listening times, and listener demographics. The software includes options for exporting data to Excel, printing directly from the tool, and receiving quarterly updates on audience metrics.

Once exported, the database from their internal tool is then exported into Excel and added to the company's SharePoint site.

3.3. Data Cleaning and Transformation

Power BI's Power Query served as a data warehouse and ETL tool for this project.

The SharePoint site provided by the client is organized into three main folders: Prebivalci (English translation: Citizens), which contains an Excel file with anonymized citizen data; Sifranti (English translation: Dimensions), containing all the dimensions in Excel files; and Vir podatkov (English translation: Data source), Excel files which include detailed information of 15-minute listening intervals of surveyed people for each day from 5:00 to 23:00.

The data from these folders was imported into Power BI's Power Query using two native connectors: Excel Workbook and SharePoint folder. One table was also created using the Enter Data option in Power BI.

Based on the tables imported from Sharepoint, we created four fact tables and six dimension tables (and one additional dimension table that was added manually). The structure of the tables can be found in Appendix B.

The tables went through data transformation such as changing data types, renaming columns for clarity, filtering, invoking custom functions, removing unnecessary columns, grouping rows, adding new columns, pivoting tables and similar.

The final tables are listed and described in the table below.

Name (Slovenian)	Name (English)	Type	Description
fact PRE Prebivalstvo	Population	fact	Shows the distribution of survey responses based on various demographic parameters such as age, gender, education level, income level, and similar.

fact PIN Poslusanost intervali	Interval Listenership	fact	It focuses on the detailed listening sessions of respondents, capturing specific intervals during which they listened to various radio stations. It includes demographic information and the exact time intervals of radio listening, along with the radio station ID and a weighting factor for each session.
fact PDN Poslusanost dnevni	Daily Listenership	fact	Represents aggregated listening data on a daily basis. It records demographic information, but the focus is on the overall daily engagement with radio stations. It includes a daily weight factor for the adjustment of representativeness.
fact PDN Poslusanost tedenski	Weekly Listenership	fact	Represents aggregated listening data on a weekly basis. It records demographic information, but the focus is on the overall weekly engagement with radio stations. It includes a weekly weight factor for the adjustment of representativeness.
dim INT Casovni intervali	Time Intervals	dim	Provides a detailed breakdown of 15-minute intervals throughout the day, ranging from 5:00 to 23:00, offering precise time segment information.
dim DOH Mesecni dohodek	Monthly Income	dim	Offers a breakdown of five income levels per month, providing financial classification information.
dim IZB Izobrazba	Education Level	dim	Categorizes various educational levels, providing straightforward details on the educational classification system.
dim TNA Tip naselja	Type of Settlement	dim	Containes information about the types of residential areas.
dim RAD Radio	Radio Station	dim	Lists all radio stations included in the research.

dim SPO Spol	Gender	dim	Categorizes gender.
dim OBC Obcine	Municipalities	dim	Lists all Slovenian municipalities along with their corresponding regions.

Table 1 – Final Tables

3.4. DATA MODELING

Calculated Tables

Throughout the data modeling phase, additional tables were created using Power BI's Data Analysis Expressions (DAX). These were not imported from external sources but generated internally within Power BI to enhance the data model. The structure of the calculated tables can be found in Appendix C.

The calculated tables are listed and described in the table below. DAX formulas used to create the tables can be found in Appendix D.

Name (Slovenian)	Name (English)	Type	Description
dim DAA Dan anketiranja	Survey Day	dim	Created as a link between *fact PTE Poslusanost tedenski, fact PIN Poslusanost intervali*, and *fact PDN Poslusanost dnevni.* It contains all days in the week on which surveys were conducted.
dim IDB ID baze	Database ID	dim	Created as a link between *fact PTE Poslusanost tedenski, fact PIN Poslusanost intervali*, and *fact PDN Poslusanost dnevni.* It consists of a list of the IDs of all databases available.
dim KOL Koledar	Calendar	dim	Necessary table for analyzing data over time. Dates span from January of the first year of the survey to December of last year of the survey.
dim STR Starost	Age	dim	Categorizes survey respondents in age groups.

Table 2 – Calculated Tables

Calculated Columns

Additionally to adding calculated tables, we also added some Calculated Columns to existing tables where needed.

The calculated columns are listed and described in the table below. DAX formulas used to create the columns can be found in Appendix D.

Table Name	Column Name (Slovenian)	Column Name (English)	Description
fact PIN Poslusanost intervali (*Interval Listenership*)	PIN Datum obdobja	Date	Created to establish a link between this table and the calendar dimension. Consists of the first days of the observed months.
fact PDN Poslusanost dnevni (*Daily Listenership*)	PDN Datum obdobja	Date	Created to establish a link between this table and the calendar dimension. Consists of the first days of the observed months.
fact PDN Poslusanost dnevni (*Daily Listenership*)	PDN Starost	Age	Created to establish a link between this table and the age dimension. Calculates the age of the respondent.
fact PTE Poslusanost tedenski (*Weekly Listenership*)	PTE Datum obdobja	Date	Created to establish a link between this table and the calendar dimension. Consists of the first days of the observed months.
fact PTE Poslusanost tedenski (*Weekly Listenership*)	PTE Starost	Age	Created to establish a link between this table and the age dimension. Calculates the age of the respondent.
dim IZB Izobrazba (*Education Level*)	IZB Izobrazba skupine	Education Level Group	Classifies individual education levels into specific groups.

Table 3 – Calculated Columns

Model and Relationships

By incorporating the newly added columns, we were able to establish relationships among the tables, and build a well-structured data model.

After relationships between tables were established, we ended up with a collection of star schemas with interlinked fact tables. This model is more widely known as a galaxy schema or a fact constellation schema. The completed schema is shown below.

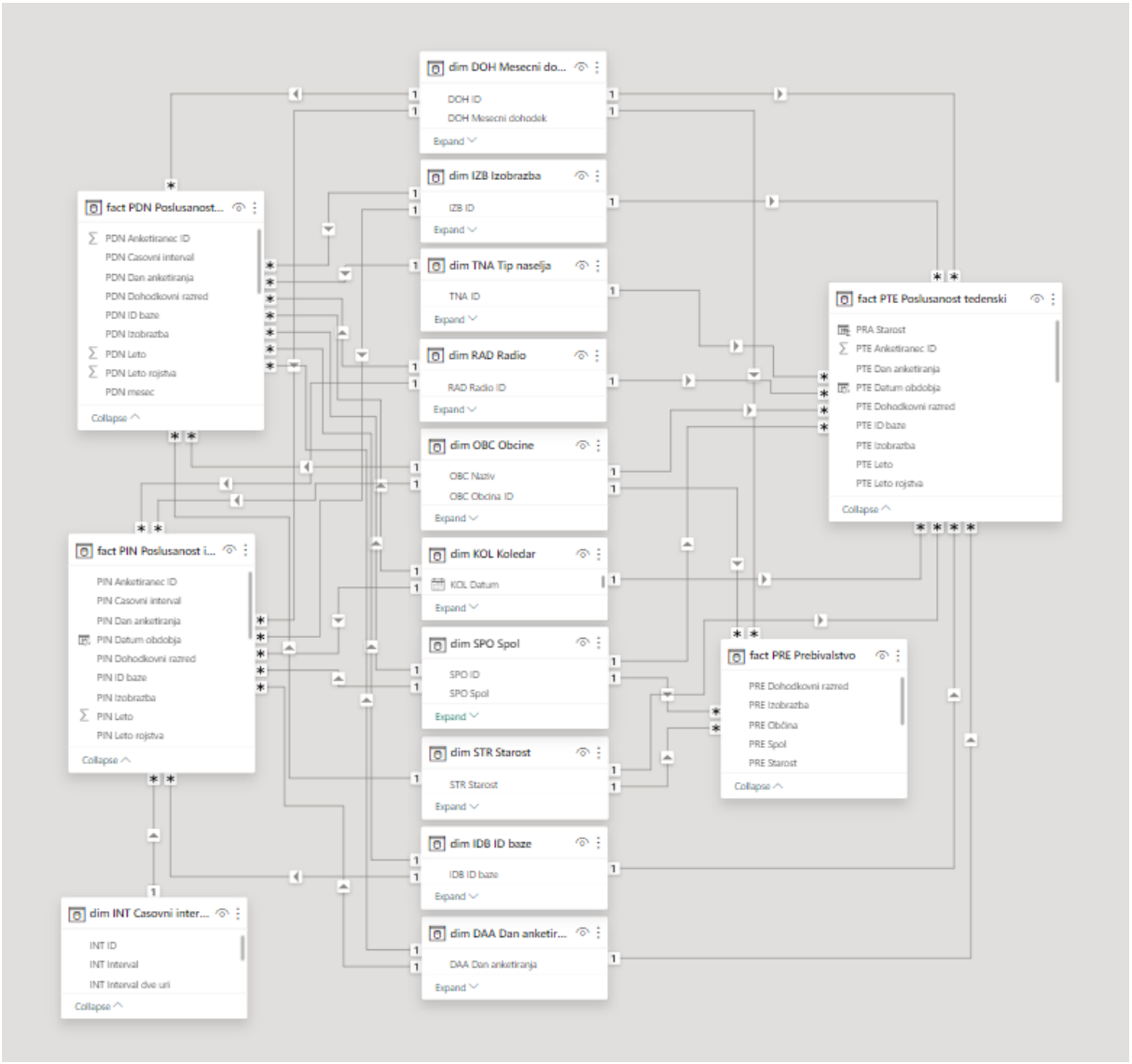


Figure 1 – Data Model

Field Parameters

Once our model was complete and all relationships established, we added two field parameters to allow end users dynamic control and change of fields in the report.

The field parameters are listed and described in the table below.

Parameter Name (Slovenian)	Parameter Name (English)	Description
Param DT	Param DT	Sets up a structure that allows users to switch between different time interval fields within the 'fact PIN Poslusanost intervali' table.
Param Povzetek	Param Summary	Sets up a structure that allows users to switch between reach in 1000 or reach in %.

Table 4 - Parameters

Measures

In the scope of Data Modeling, we also created some measures using DAX formulas. The measures supported our analysis, allowed us to perform dynamic calculations across our data sets, and enhanced the data exploration experience.

The list of our measures can be found in Appendix D.

3.5. DATA VISUALIZATION

After setting up our data model, we were able to focus on the visualization phase of the project. This step transformed our data into clear visuals that assisted in decision-making. Our aim was to tell a story with data that was easy for every one of our client's users to understand and use.

We also focused on making the report interactive. Users could look at data in different ways, focus on certain details, or dig deeper into the numbers. This made our report a dynamic tool for discovering trends, spotting opportunities, or finding areas that needed attention.

By moving from organizing our data to building this report, we hoped to turn our complex data into simple, insightful visuals.

Our report ended up consisting of five pages, with one of the pages hidden and serving as a drill-through page.

The pages were the following:

- Dnevni doseg (English translation: Daily Reach)
- Tedenski doseg (English translation: Weekly Reach)
- Intervali (English translation: Intervals)
- Povzetek (English translation: Summary)
- DT:radio

Every page featured a custom filter pane positioned on the left side, enabling users to filter the data according to demographic attributes such as age, gender, region, municipality, monthly income, and education level. Additionally, this pane allowed filtering by time period, particular radio station, and day of the week, offering users a flexible tool to drill down into the dataset for more specific insights. The filter pane consisted of main filters which were always show, and additional filters that could be accessed by clicking on the Več filtrov + (English translation: More filters +). It also included a feature to reset all filters in the bottom of the pane Ponastavi filtre (Reset filters). These functions were all accomplished with the use of bookmarks.

All pages, with the exception of the drill-through page, also incorporated page navigation buttons in the bottom right corner to allow easier transition between pages.

Daily Reach and Weekly Reach (Dnevni doseg in Tedenski doseg)

These two pages included identical visuals but looked at from different perspectives: one for daily reach and the other for weekly reach. The daily reach was calculated as an average across all seven days covered in the survey, while the weekly reach summed up the totals for these days.

As detailed in Section 3.1 of this thesis, the survey sampled 2,016 Slovenian residents aged between 15 and 80 years, representing a carefully chosen sample of the population, referred to as the numerus in the Power BI report. The data collected from these respondents gave us a percentage of the population that listened to specific radio stations either daily or weekly. This percentage was then extrapolated to represent the entire Slovenian population within the specified age group. Consequently, we obtained the actual reach (expressed here in thousands) for this demographic.

The matrix table located at the upper left of both pages displayed this information, listing each radio station's name alongside its reach percentage and the reach figure extrapolated to thousands. Additionally, both pages provided access to the drill-through page via the drill-through option available on the radio station column of the table, which could be accessed through options that appear upon right-clicking with the mouse.

The two pages also included two additional visuals: a donut chart showing the percentage distribution of data by gender, and a clustered bar chart displaying the reach in thousands according to the regions of Slovenia where the respondents lived.

Besides the above-mentioned filter pane and page navigation buttons these two pages also included a box with information regarding the period selected, the numerus size, total reach in percentage and reach in thousands. This information was dynamically adjusted based on the filters selected in the filter pane.

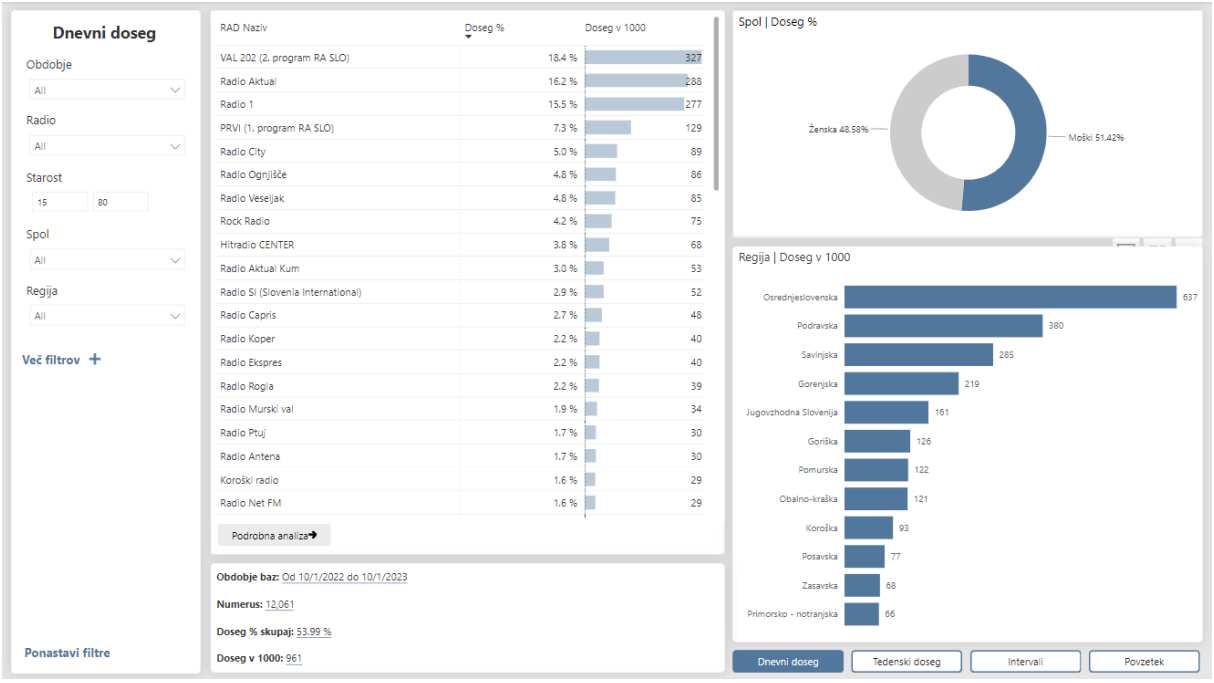


Figure 2 – Daily Reach Page

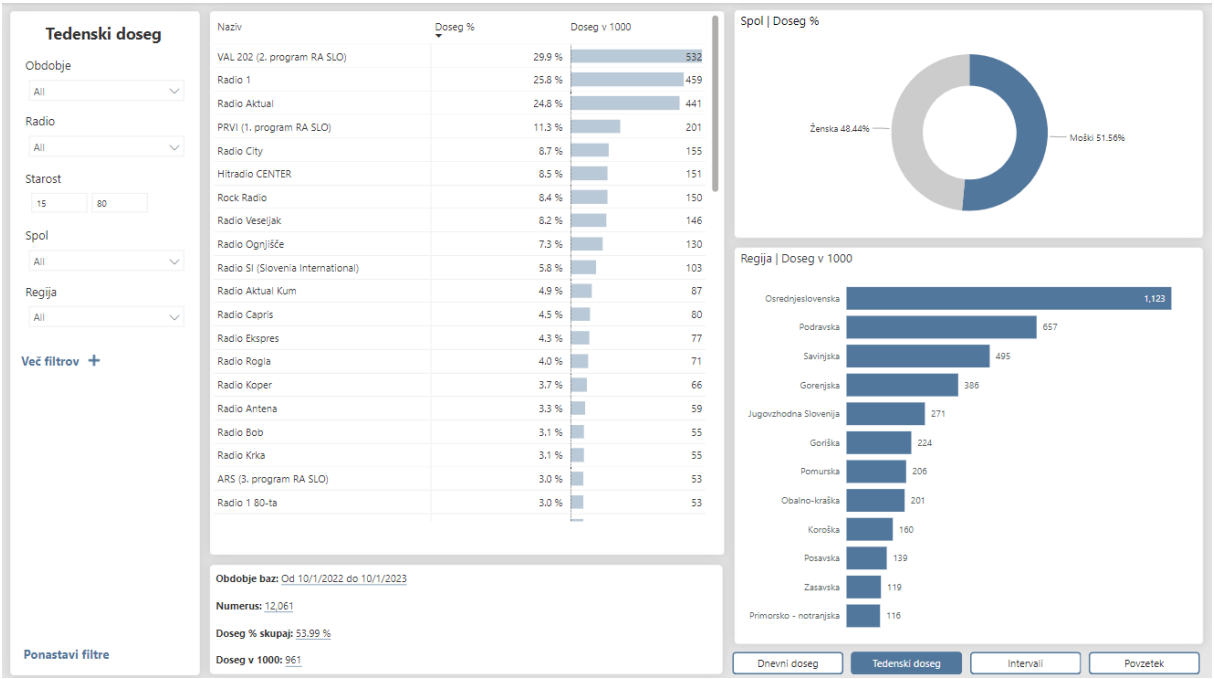


Figure 3 – Weekly Reach Page

Intervals (Intervali)

Beside the filter pane and the page navigation buttons, the majority of this page consisted of two matrix tables that the users could switch between using a button at the top. This functionality was enabled with the use of bookmarks.

The default matrix was named Radio and showed a breakdown of listenership percentages by specific time intervals throughout the day. Each row represented a specific 15-minute interval of the day between 5:00 to 23:00, and each column corresponded to a different radio station. The percentages represent the share of listeners for each radio station within the given time slot.

The second option was a matrix accessed by bookmark called Radio/Dan (English translation: Radio/Day). This matrix also showed a breakdown of listenership percentages by specific time intervals throughout the day. Furthermore, this option also included a breakdown by days of the week, with columns for each day from Monday to Sunday. This allowed for analysis of listenership patterns not just throughout the day but also across the different days of the week.

Similarly to the Daily Reach and Weekly Reach, this table also had a box with information regarding the numerus size.

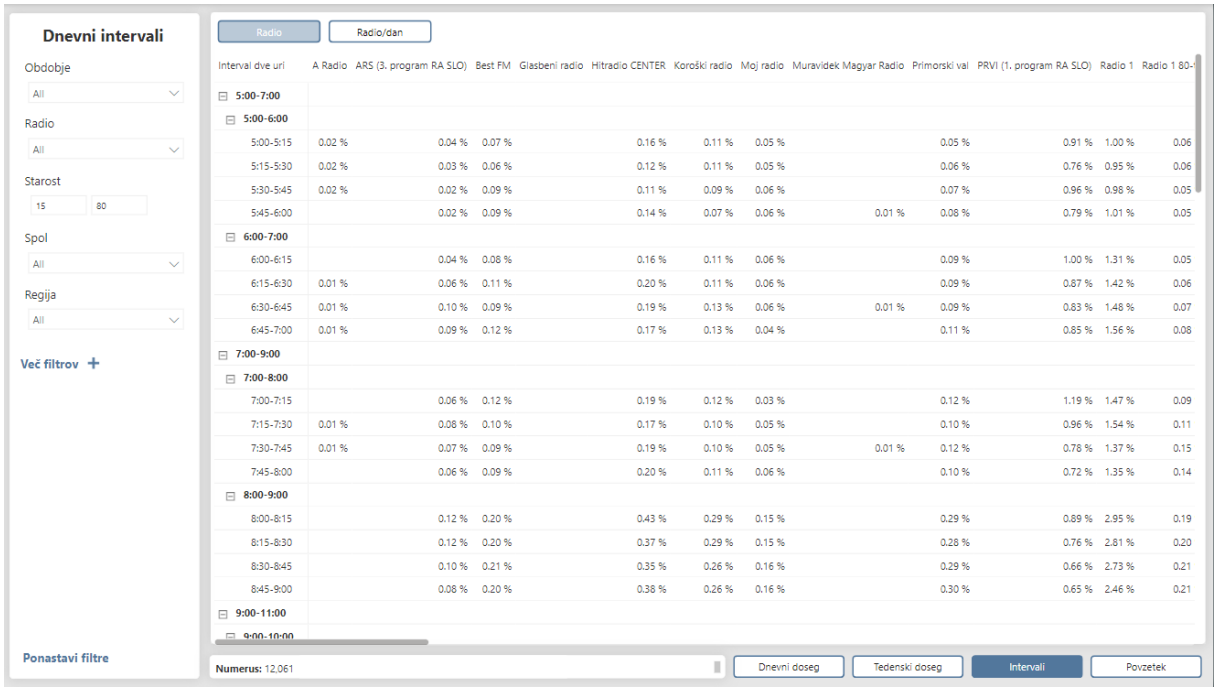


Figure 4 – Intervals (Radio View)

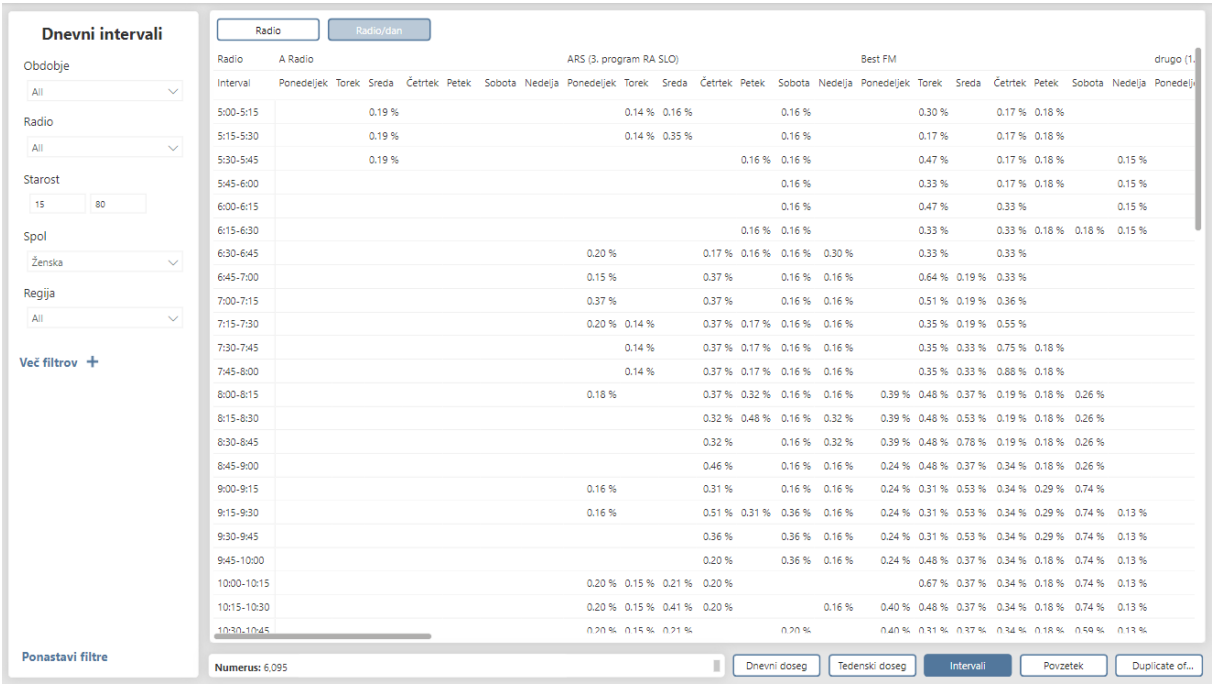


Figure 5 – Intervals (Radio/dan View)

Summary (Povzetek)

This page serves as a summary.

In addition to the custom filter pane, the page navigation buttons, and a dedicated information box detailing the numerus size, the page consisted of three distinct visuals.

On the upper left side was a donut chart that outlined the audience percentage reach by gender. On the right side, we positioned a clustered bar chart that showed the reach in thousands, segmented by the regions of Slovenia where the survey participants were based. Below these two charts is a clustered column chart. This chart presented the audience reach, offering a toggle between thousands and percentages for individual radio stations. This functionality was achieved through a switch enabled by a pair of buttons, using the parameter 'Param Povzetek' (Param Summary).

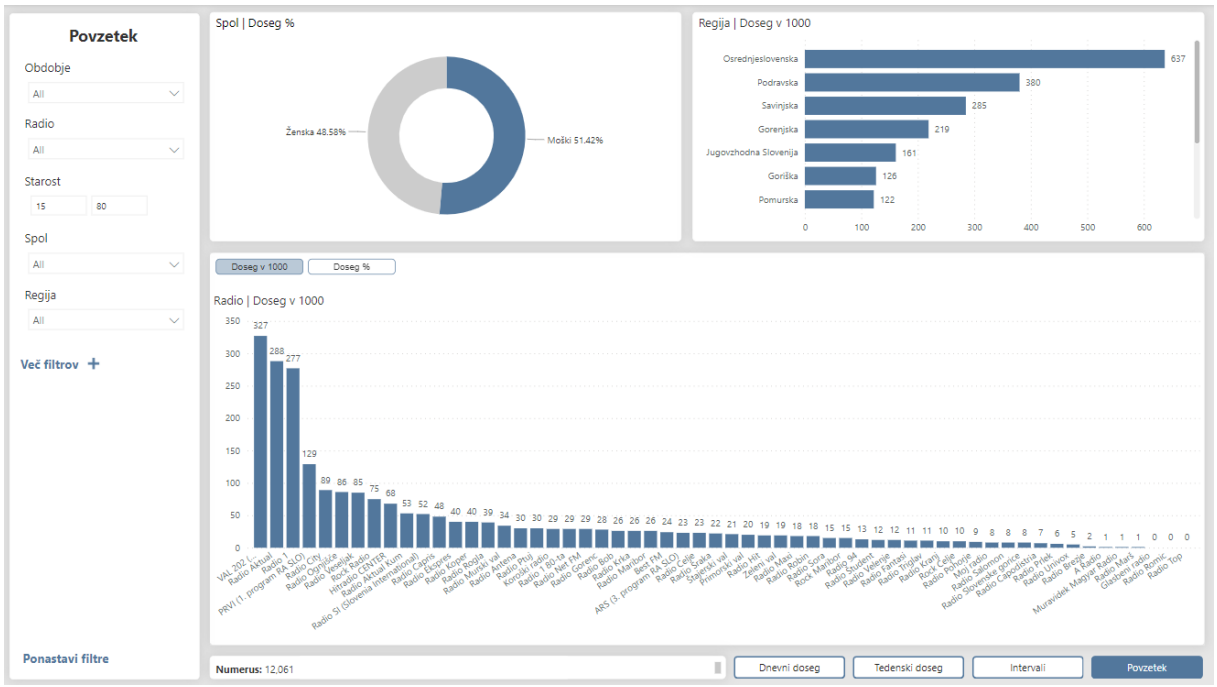


Figure 6 – Summary (Reach in 1000)

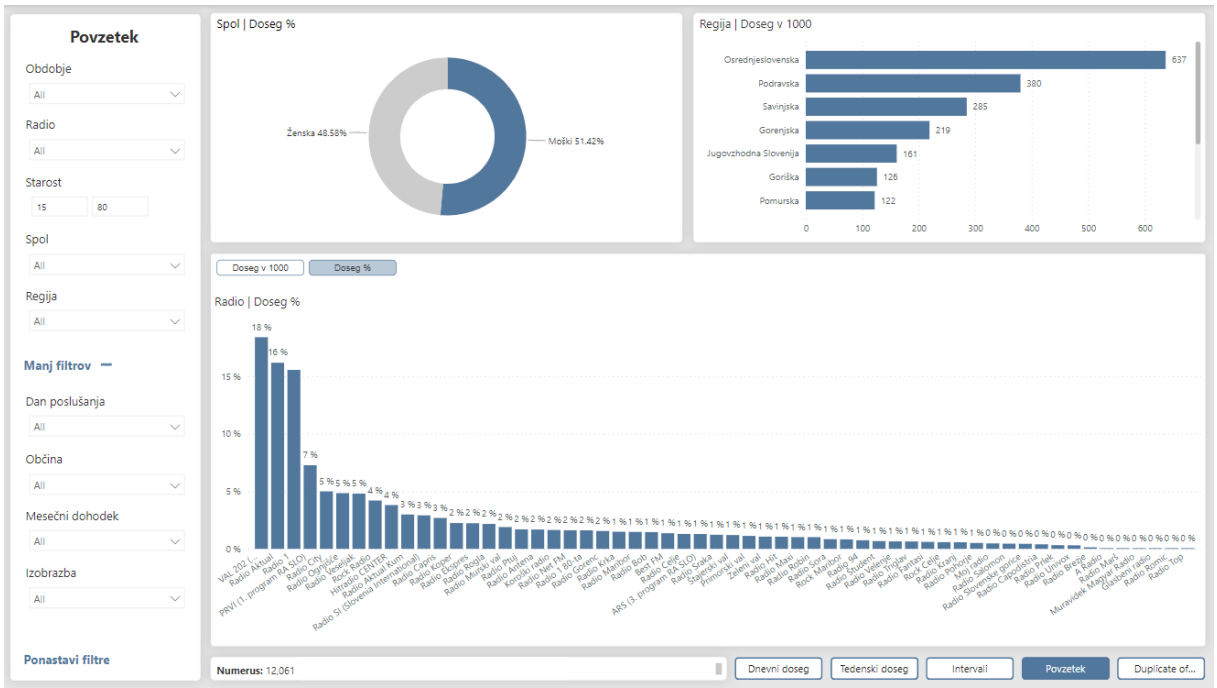


Figure 7 – Summary (Reach %)

Drill-Through Page (DT:radio)

As previously mentioned, this page could be navigated to by a drill-through action on a particular radio station from any of the previous pages. The page provided an in-depth analysis of the selected radio station's listenership data.

The page included a header that clearly indicated the radio station being analyzed. It displayed a donut chart that split reach in percentages by gender. Additionally, it consisted of four matrix tables, each detailing the station's reach in terms of percentages and in thousands, categorized by specific demographics such as age group, educational background, monthly income, and the geographical location of the survey participants.

Moreover, the page included a line chart that showed the station's percentage reach throughout various intervals over a day. This chart contains three lines correlating to different listener categories. These categories differentiated the data by the entire week, weekends only, and weekdays only. Users had the option to select which category lines to display on the chart, a feature enabled by the Param DT parameter.

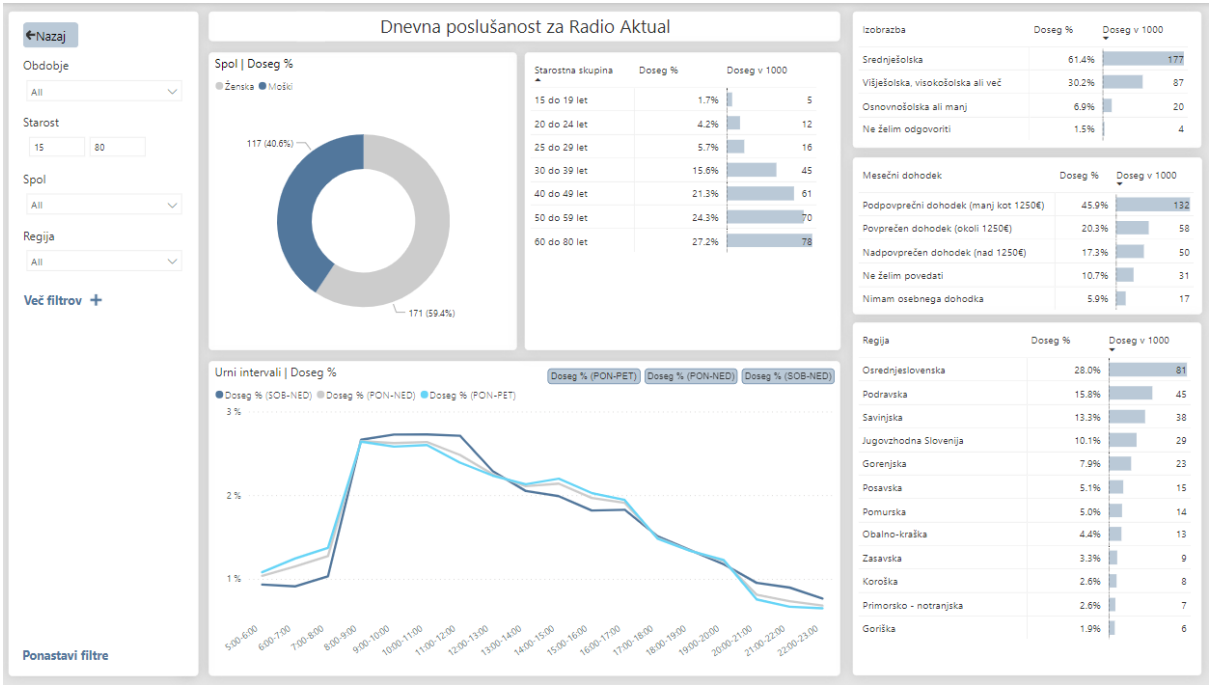


Figure 8 – Drill-Through Page

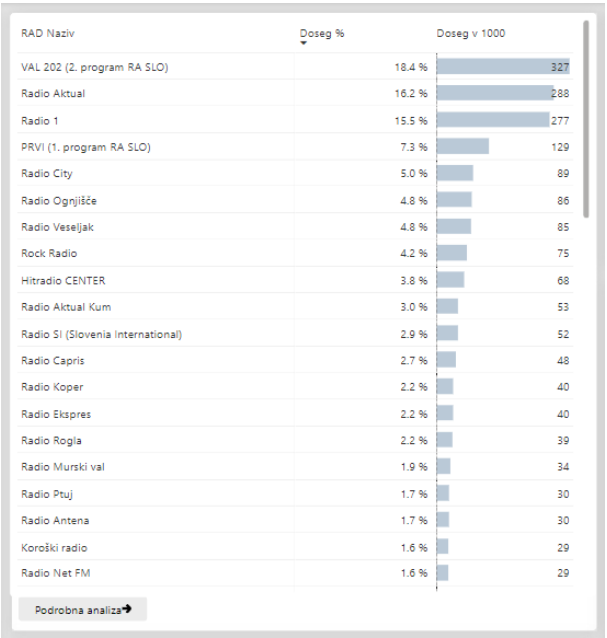
4. RESULTS AND DISCUSSION

4.1. RESULTS

For the purpose of interpreting results, I provided some deeper analysis. Therefore, some of the visuals below cannot be seen in the report we delivered to the client.

Top Radio – VAL 202

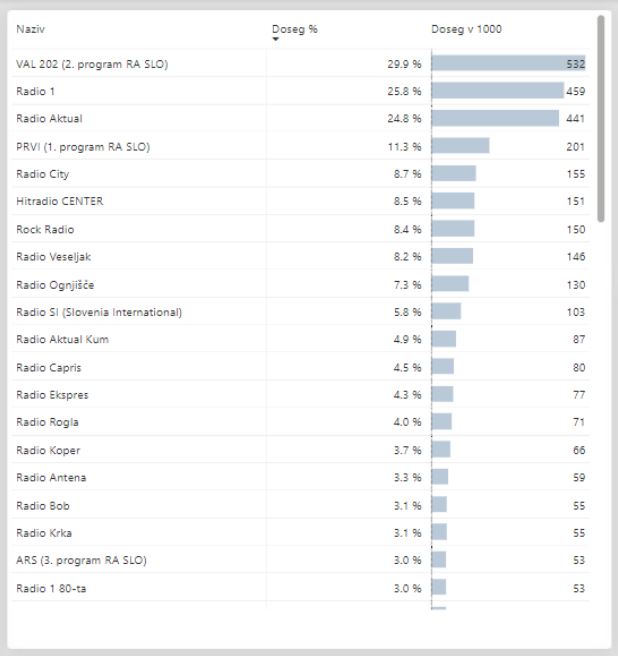
Without any filters applied and therefore using the entire dataset, VAL 202 (2. program RA SLO) appeared to be the most popular radio station. It achieved a daily reach of 327 thousand individuals within the 15-80 year age bracket throughout Slovenia, accounting for 18.4% of the population in this demographic. On a weekly basis, the station's reach extended to an average of 532 thousand people, or 29.9% of the population.



RAD Naziv	Doseg %	Doseg v 1000
VAL 202 (2. program RA SLO)	18.4 %	327
Radio Aktual	16.2 %	288
Radio 1	15.5 %	277
PRVI (1. program RA SLO)	7.3 %	129
Radio City	5.0 %	89
Radio Ognjišče	4.8 %	86
Radio Veseljak	4.8 %	85
Rock Radio	4.2 %	75
Hitradio CENTER	3.8 %	68
Radio Aktual Kum	3.0 %	53
Radio SI (Slovenia International)	2.9 %	52
Radio Capris	2.7 %	48
Radio Koper	2.2 %	40
Radio Ekspres	2.2 %	40
Radio Rogla	2.2 %	39
Radio Murski val	1.9 %	34
Radio Ptuj	1.7 %	30
Radio Antena	1.7 %	30
Koroški radio	1.6 %	29
Radio Net FM	1.6 %	29

Podrobna analiza➔

Figure 9 – Daily Reach

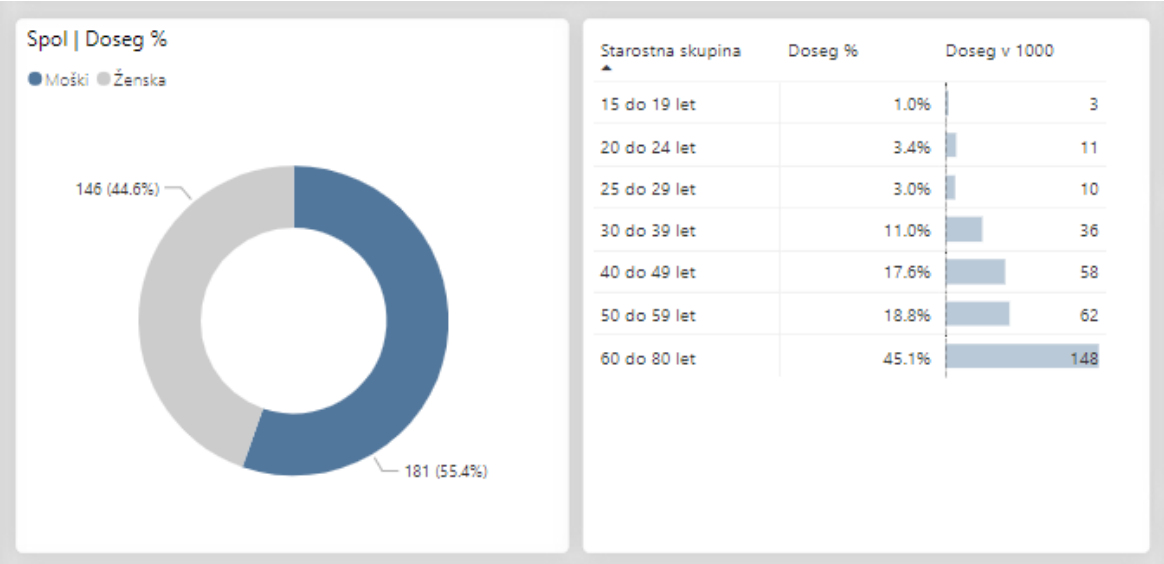


Naziv	Doseg %	Doseg v 1000
VAL 202 (2. program RA SLO)	29.9 %	532
Radio 1	25.8 %	459
Radio Aktual	24.8 %	441
PRVI (1. program RA SLO)	11.3 %	201
Radio City	8.7 %	155
Hitradio CENTER	8.5 %	151
Rock Radio	8.4 %	150
Radio Veseljak	8.2 %	146
Radio Ognjišče	7.3 %	130
Radio SI (Slovenia International)	5.8 %	103
Radio Aktual Kum	4.9 %	87
Radio Capris	4.5 %	80
Radio Ekspres	4.3 %	77
Radio Rogla	4.0 %	71
Radio Koper	3.7 %	66
Radio Antena	3.3 %	59
Radio Bob	3.1 %	55
Radio Krka	3.1 %	55
ARS (3. program RA SLO)	3.0 %	53
Radio 1 80-ta	3.0 %	53

Figure 10 – Weekly Reach

The prime listening time for VAL 202 was observed to be from 7:00 to 11:00.

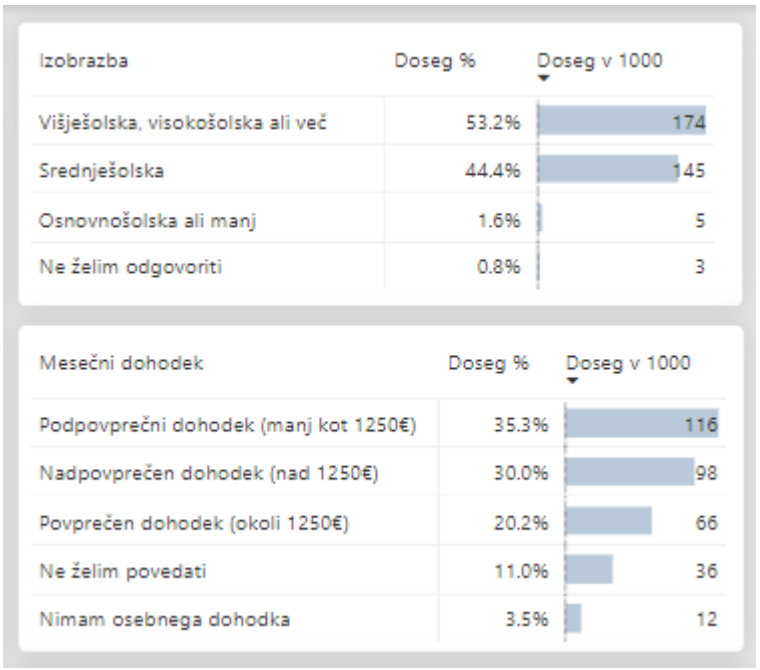
Regarding demographics, the listenership mainly consisted of male audiences, who made up 55.4%, with female listeners following closely at 44.6%. The 60-80 year age group comprised the largest segment of VAL 202's audience, with a 45.1% share.



Starostna skupina	Doseg %	Doseg v 1000
15 do 19 let	1.0%	3
20 do 24 let	3.4%	11
25 do 29 let	3.0%	10
30 do 39 let	11.0%	36
40 do 49 let	17.6%	58
50 do 59 let	18.8%	62
60 do 80 let	45.1%	148

Figure 11 – Reach of VAL 202 by Gender and Age

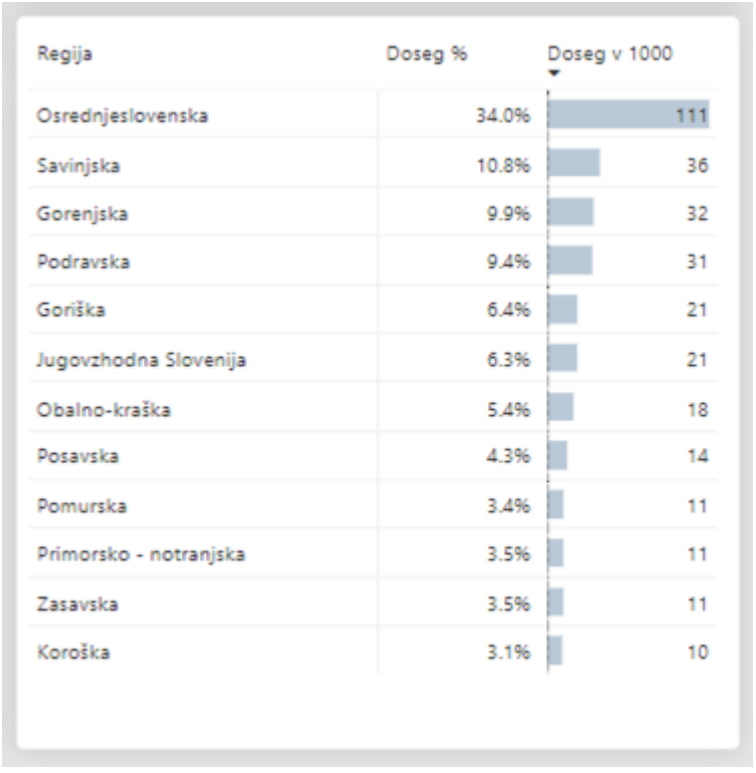
The majority of VAL 202's listeners were highly educated, comprising 53.2% of its audience, and a big portion, 35.3%, had an above-average monthly income. Geographically, a significant number of these listeners, specifically 34%, resided in the Central Slovenia region.



Izobrazba	Doseg %	Doseg v 1000
Višješolska, visokošolska ali več	53.2%	174
Srednješolska	44.4%	145
Osnovnošolska ali manj	1.6%	5
Ne želim odgovoriti	0.8%	3

Mesečni dohodek	Doseg %	Doseg v 1000
Podpovprečni dohodek (manj kot 1250€)	35.3%	116
Nadpovprečen dohodek (nad 1250€)	30.0%	98
Povprečen dohodek (okoli 1250€)	20.2%	66
Ne želim povedati	11.0%	36
Nimam osebnega dohodka	3.5%	12

Figure 12 – Reach of VAL 202 by Education and Monthly Income Reach

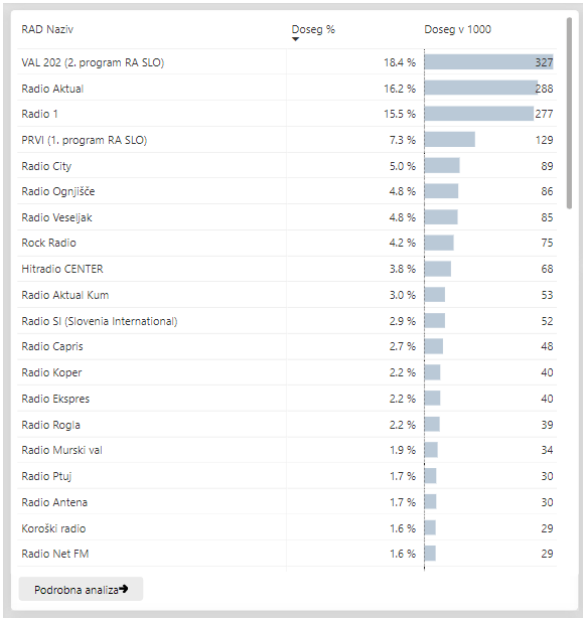


Regija	Doseg %	Doseg v 1000
Osrednjeslovenska	34.0%	111
Savinjska	10.8%	36
Gorenjska	9.9%	32
Podravska	9.4%	31
Goriška	6.4%	21
Jugovzhodna Slovenija	6.3%	21
Obalno-kraška	5.4%	18
Posavska	4.3%	14
Pomurska	3.4%	11
Primorsko - notranjska	3.5%	11
Zasavska	3.5%	11
Koroška	3.1%	10

Figure 133 – Reach of VAL 202 by Region Reach

General Results

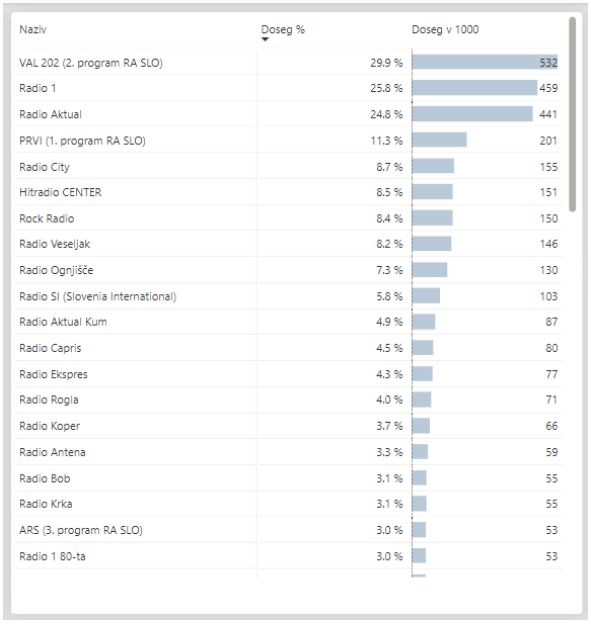
Overall, the most widely favored radio stations, both on a daily and weekly basis, included VAL 202 (with a daily reach of 18.4% and a weekly reach of 29.9%), Radio Aktual (with a daily reach of 16.2% and a weekly reach of 24.8%), Radio 1 (with a daily reach of 15.5% and a weekly reach of 25.8%), PRVI (with a daily reach of 7.3% and a weekly reach of 11.3%), and Radio City (with a daily reach of 5.0% and a weekly reach of 8.7%).



RAD Naziv	Doseg %	Doseg v 1000
VAL 202 (2. program RA SLO)	18.4 %	327
Radio Aktual	16.2 %	288
Radio 1	15.5 %	277
PRVI (1. program RA SLO)	7.3 %	129
Radio City	5.0 %	89
Radio Ognjišče	4.8 %	86
Radio Veseljak	4.8 %	85
Rock Radio	4.2 %	75
Hitradio CENTER	3.8 %	68
Radio Aktual Kum	3.0 %	53
Radio SI (Slovenia International)	2.9 %	52
Radio Capris	2.7 %	48
Radio Koper	2.2 %	40
Radio Ekspres	2.2 %	40
Radio Rogla	2.2 %	39
Radio Murski val	1.9 %	34
Radio Ptuj	1.7 %	30
Radio Antena	1.7 %	30
Koroški radio	1.6 %	29
Radio Net FM	1.6 %	29

Podrobna analiza

Figure 14 – Daily reach



Naziv	Doseg %	Doseg v 1000
VAL 202 (2. program RA SLO)	29.9 %	532
Radio 1	25.8 %	459
Radio Aktual	24.8 %	441
PRVI (1. program RA SLO)	11.3 %	201
Radio City	8.7 %	155
Hitradio CENTER	8.5 %	151
Rock Radio	8.4 %	150
Radio Veseljak	8.2 %	146
Radio Ognjišče	7.3 %	130
Radio SI (Slovenia International)	5.8 %	103
Radio Aktual Kum	4.9 %	87
Radio Capris	4.5 %	80
Radio Ekspres	4.3 %	77
Radio Rogla	4.0 %	71
Radio Koper	3.7 %	66
Radio Antena	3.3 %	59
Radio Bob	3.1 %	55
Radio Krka	3.1 %	55
ARS (3. program RA SLO)	3.0 %	53
Radio 1 80-ta	3.0 %	53

Figure 15 – Weekly Reach

In general, the majority of listeners were men at 51.42%, with women representing 48.58%. Most of the listeners resided in the Central Slovenia region (637 thousand). The next regions with the most listeners were Drava (380 thousand), Savinja (285 thousand) and Upper Carniola (219 thousand).

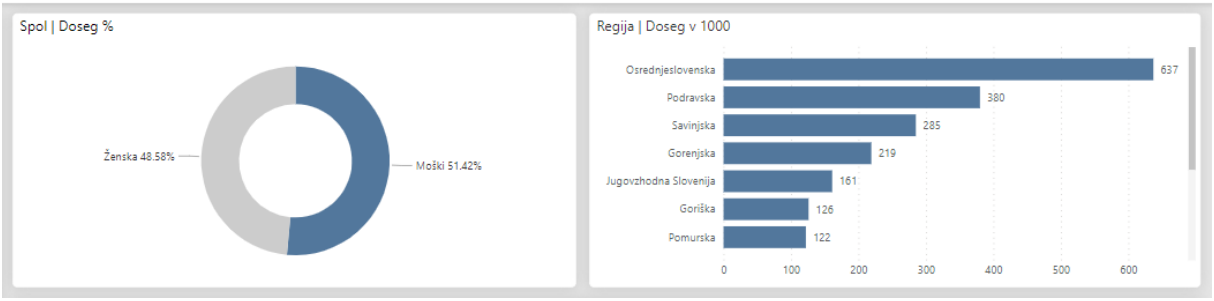


Figure 16 – Reach by Gender and Region Reach

Among women, Radio Aktual (with a daily reach of 19.0% and a weekly reach of 28.3%) and Radio 1 (with a daily reach of 17.1% and a weekly reach of 27.6%) demonstrated higher popularity compared to VAL 202 (with a daily reach of 16.2% and a weekly reach of 26.6%), which was the most popular radio station across the entire population.

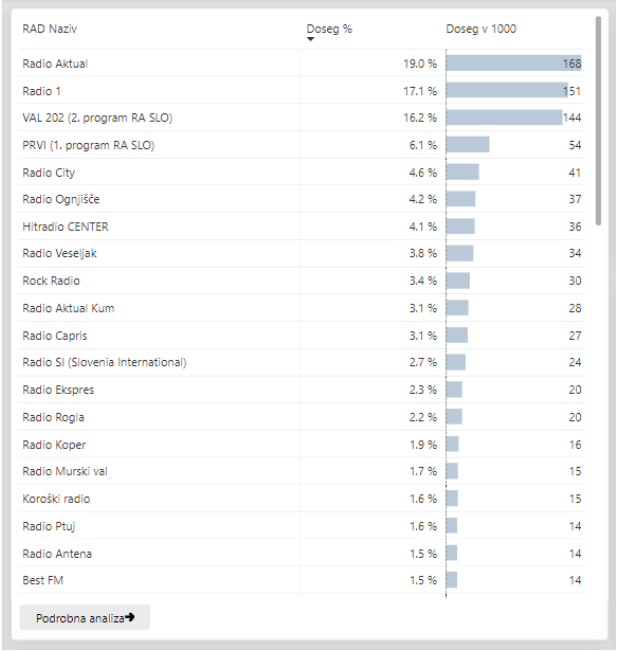
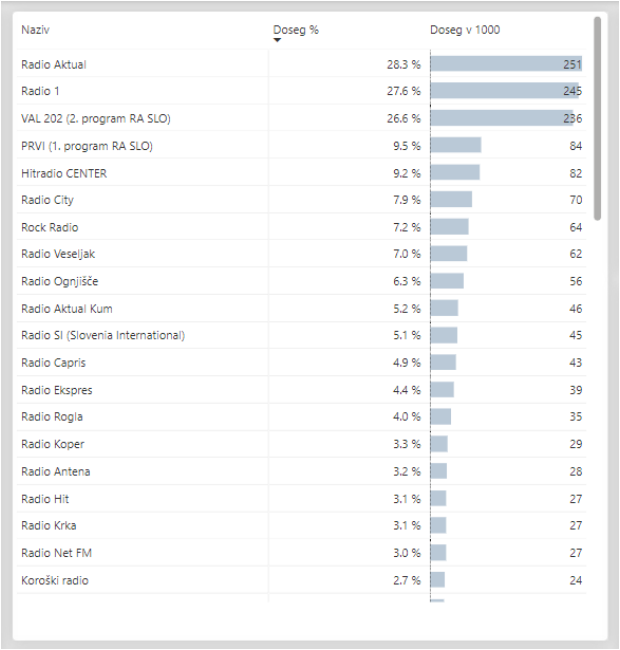


Figure 17 – Daily Reach Women

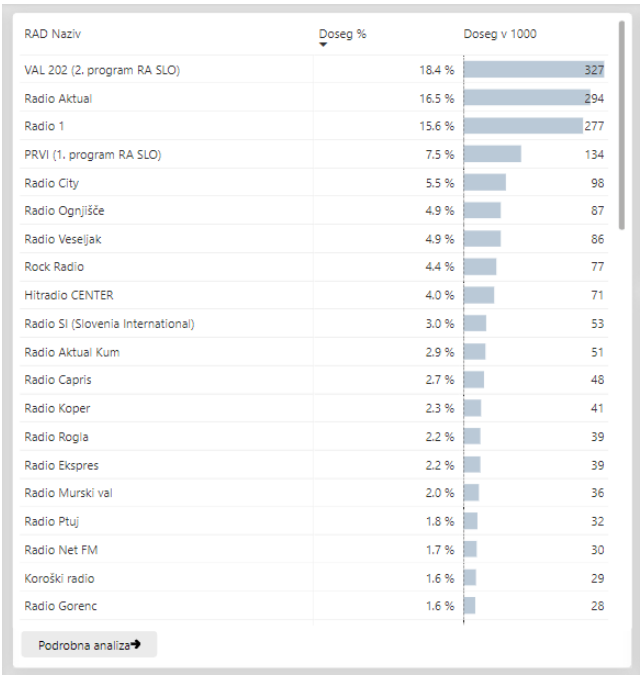


Naziv	Doseg %	Doseg v 1000
Radio Aktual	28.3 %	251
Radio 1	27.6 %	245
VAL 202 (2. program RA SLO)	26.6 %	236
PRVI (1. program RA SLO)	9.5 %	84
Hitradio CENTER	9.2 %	82
Radio City	7.9 %	70
Rock Radio	7.2 %	64
Radio Veseljak	7.0 %	62
Radio Ognjišče	6.3 %	56
Radio Aktual Kum	5.2 %	46
Radio SI (Slovenia International)	5.1 %	45
Radio Capris	4.9 %	43
Radio Ekspres	4.4 %	39
Radio Rogla	4.0 %	35
Radio Koper	3.3 %	29
Radio Antena	3.2 %	28
Radio Hit	3.1 %	27
Radio Krka	3.1 %	27
Radio Net FM	3.0 %	27
Koroški radio	2.7 %	24

Figure 18 – Weekly Reach Women

Aside from this, listening habits among both genders followed a consistent pattern.

During the earlier observed periods, spanning October 2022, January 2023, April 2023, and May 2023, Radio Aktual emerged as the second most popular station. However, a shift was noted in the later months of 2023, specifically in July and October, when Radio 1 surpassed Radio Aktual.



RAD Naziv	Doseg %	Doseg v 1000
VAL 202 (2. program RA SLO)	18.4 %	327
Radio Aktual	16.5 %	294
Radio 1	15.6 %	277
PRVI (1. program RA SLO)	7.5 %	134
Radio City	5.5 %	98
Radio Ognjišče	4.9 %	87
Radio Veseljak	4.9 %	86
Rock Radio	4.4 %	77
Hitradio CENTER	4.0 %	71
Radio SI (Slovenia International)	3.0 %	53
Radio Aktual Kum	2.9 %	51
Radio Capris	2.7 %	48
Radio Koper	2.3 %	41
Radio Rogla	2.2 %	39
Radio Ekspres	2.2 %	39
Radio Murski val	2.0 %	36
Radio Ptuj	1.8 %	32
Radio Net FM	1.7 %	30
Koroški radio	1.6 %	29
Radio Gorenc	1.6 %	28

Podrobna analiza

Figure 19 – Daily Reach (October 2022, January 2023, April 2023, and May 2023)

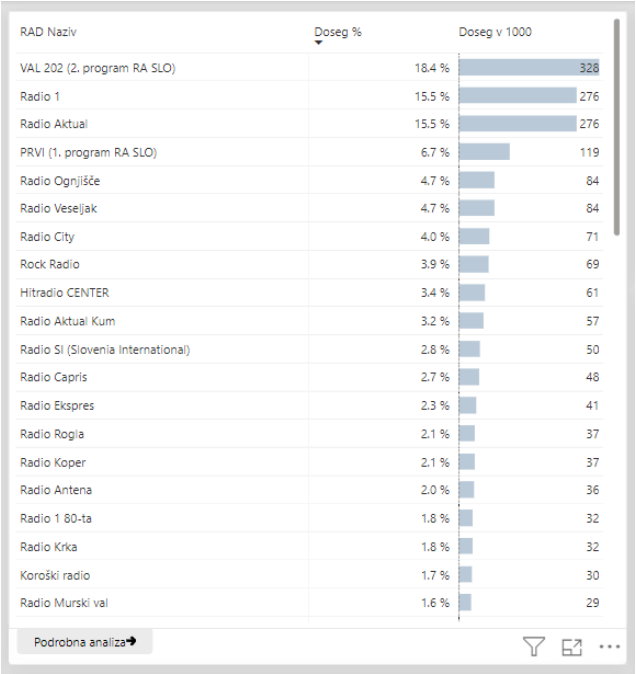


Figure 20 – Daily Reach (July and October 2023)

The highest reach occurs in the middle of the week, with Tuesday and Thursday both showing a reach of 20%. Monday, Wednesday, and Friday all show 19% reach, which is only slightly lower than the peak days. There's a noticeable drop during the weekend, with Saturday at 18% and Sunday at 17%. While the decrease is not drastic, it suggests that the audience may engage in different activities during the weekend that compete with radio listening.

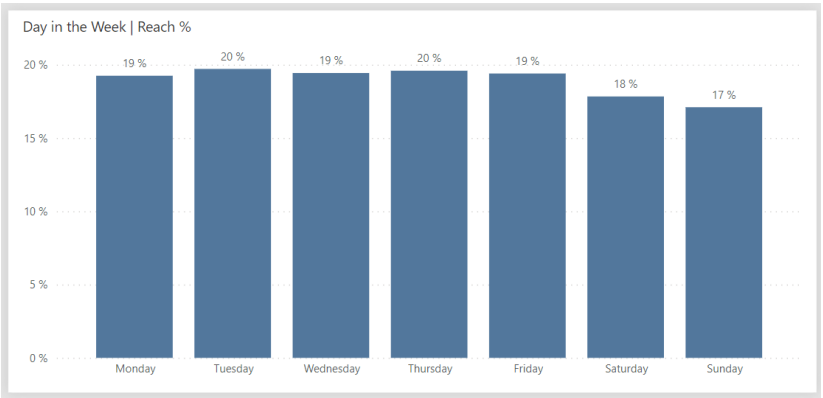


Figure 21 – Reach by Day in the Week

There is a significant spike in listenership during the morning hours. The highest reach is at 8:00-9:00 with 20%, closely followed by 9:00-10:00 with 19%. This pattern suggests that many listeners tune in to the radio during their morning routine or commute. Between 10:00-14:00, the listenership remains relatively high, ranging from 15% to 18%. This could indicate that the

radio is a popular choice for background listening during work hours or that people listen during their midday activities. After 17:00, there is a noticeable drop in listenership, with 9%, hitting the lowest points in the late evening and night, with 3% from 20:00 onwards.

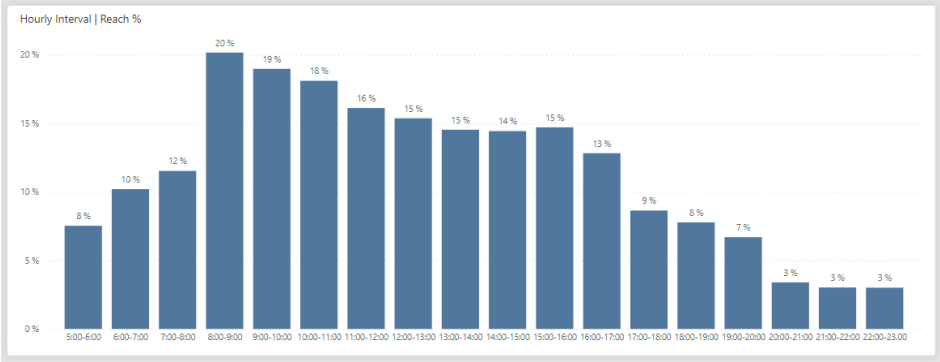


Figure 22 – Reach by Hourly Interval

The largest reach was among individuals with secondary education, with a reach of 1,258 thousand individuals. The next significant post-secondary education (college, university, or higher), with a reach of 949 thousand. This suggests a strong engagement from those with higher education, though not as high as the secondary education group. Elementary education or less, has a much smaller reach of 116 thousand. This category has the lowest reach among those with formal education levels represented in the graph. The category "I do not wish to answer" shows that 33 thousand individuals chose not to disclose their education level.

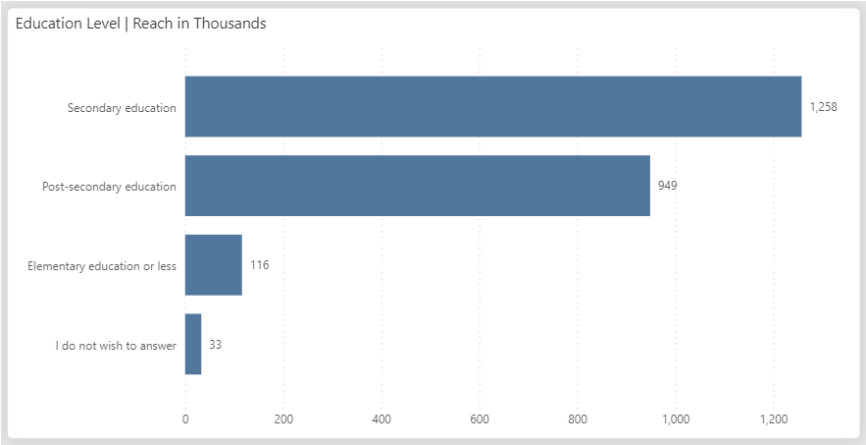


Figure 23 – Reach by Education Level

The largest reach is among individuals with an income below the average (less than 1250€), with a reach of 1,005 thousand individuals. Individuals with an income above the average (more than 1250€) represent the second-largest audience segment, with a reach of 517 thousand. This demonstrates that the stations also have a significant appeal to listeners with

higher incomes, though less than half compared to the lower income group. The reach among individuals with an average income (around 1250€) is 438 thousand, placing this segment third in terms of radio reach. A notable number of individuals, 259 thousand, prefer not to disclose their income. The reach in the last category is 136 thousand, indicating that the stations also reach individuals who may not have a personal income, such as students, retirees, or unemployed individuals.

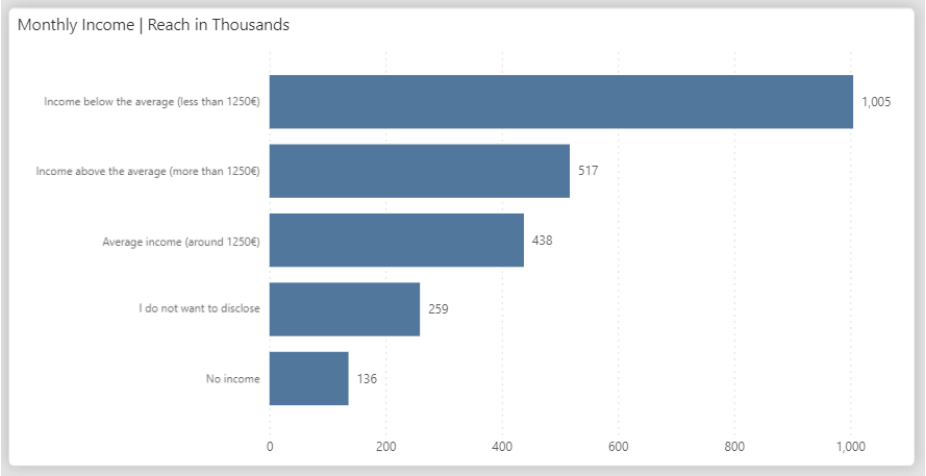


Figure 24 – Reach by Monthly Income

Leading the graph with a reach of 49 thousand, Ljubljana has the highest listenership. This is expected as Ljubljana is the capital and the largest city in Slovenia, which would naturally have a higher number of listeners due to its larger population. The second-highest reach is in Maribor, with 7 thousand listeners. Maribor is the second-largest city in Slovenia, and although the reach is significantly lower than in Ljubljana, it still represents the second-largest audience among the municipalities listed. The city of Kranj shows a reach of 2 thousand, placing it third in terms of reach. Several other municipalities such as Celje, Domžale, Novo Mesto, and Velenje each have a reach of 1 thousand listeners. The presence of listeners in these areas indicates a dispersed listenership across the country.

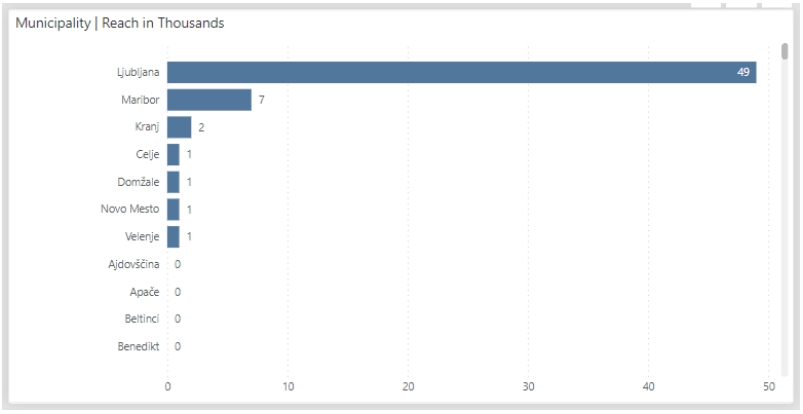


Figure 25 – Reach by Municipality

Despite the results of the previous graph, the reach in rural areas is the highest, with 1,207 thousand listeners. This suggests that the radio station or service has a strong presence in the countryside or non-urban areas. Urban areas have a slightly lower reach, with 1,132 thousand listeners. A small number of individuals, 16 thousand, either do not know or do not wish to disclose their type of settlement.

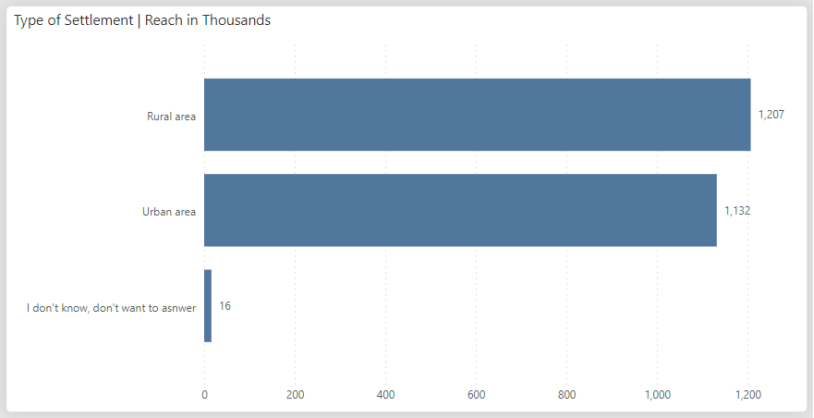


Figure 26 – Reach by Type of Settlement

The 60 to 80 years age group has the highest reach at 819 thousand. This suggests that the service is most popular or most utilized among the older population. The next highest reach is in the 50 to 59 years age group, with 453 thousand. Following closely is the 40 to 49 years age group, with a reach of 431 thousand. There's a noticeable drop for the 30 to 39 years age group, with a reach of 333 thousand, suggesting lower engagement among this demographic. The reach further decreases for the 20 to 24 years age group, with 113 thousand. Similarly, the 25 to 29 years age group has a reach of 109 thousand, indicating that the service is less popular with younger adults. The lowest reach is in the youngest age group, 15 to 19 years, with 32 thousand.

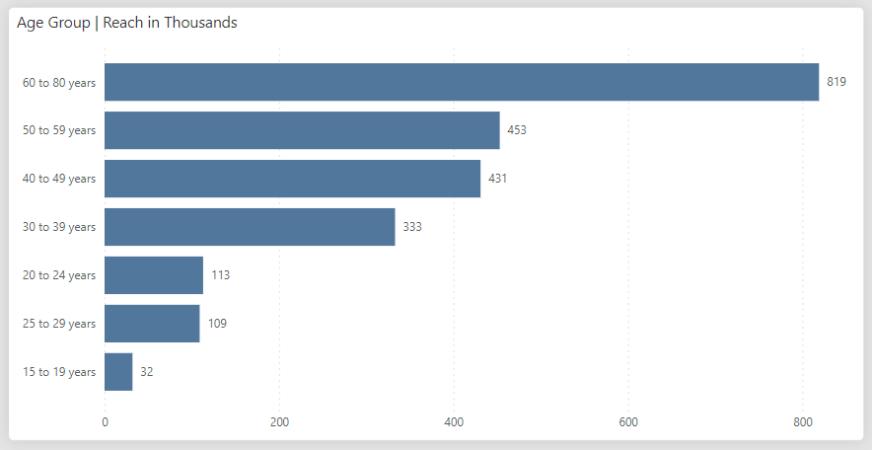


Figure 27 – Reach by Age Group

4.2. DISCUSSION

This section of the thesis aims to frame the research findings within the broader scope of radio broadcasting, specifically within the Slovenian market.

Radio Station Popularity and Reach

The results of this thesis create a hierarchy of popularity among Slovenian radio stations. With VAL 202 being the most popular radio station in both daily and weekly reach, and other stations such as Radio Aktual, Radio 1, PRVI, and Radio City following closely behind. This spread of listenership highlights a diverse radio market with varying content preferences across the population.

Demographic Distribution of Listeners

The gender distribution of radio listeners suggests that radio listening is more or less evenly distributed among genders, with the reach among men only slightly higher than females. Based on these insights, radio stations could use some subtle programming strategies to attract more female listeners.

The age distribution of the audience showed a strong presence among the older population, particularly those between 60 to 80 years old. This could be due to the established habit of radio listening within this demographic or a preference for traditional media over digital alternatives. On the other hand, younger age groups had lower engagement which indicated that radio faces competition from other media, presumably digital platforms. This generational shift signals a potential area for radio broadcasters to innovate and adapt their content to attract younger listeners.

The higher engagement of listeners with secondary and higher education could indicate that radio content is tailored towards or more accessible to these educational groups. The income data suggests that radio remains a popular medium across various income levels, with a notable reach among those below the average income. This might influence the type of advertising and sponsorship that radio stations seek, aiming to match the consumer power of their audience.

Geographically, the highest concentration of listeners is in the urban centers of Ljubljana and Maribor, which aligns with the two being the largest cities in Slovenia. The reach within smaller municipalities shows a scattered audience across the country, which indicates the importance of local content. The reach in rural areas surpasses that in urban areas, suggesting that radio plays an important role in non-urban areas, possibly due to fewer alternative entertainment options or a stronger tradition of radio listening.

Implications for Radio Broadcasting Business

These insights have significant implications for the radio broadcasting business. Stations can refine their content strategy based on demographic preferences, optimize ad placement

according to listening patterns, and explore new formats to engage younger audiences, who are currently less represented among listeners.

Limitations

As all research, this project was also limited by its chosen methods and the conditions under which the research was conducted. Recognizing and exploring these limitations is an important part of honest evaluation of the outcomes of the research. Not only that, but it also sets the foundation for any expansion of the analysis.

1. **Data Scope:** The current analysis is limited to demographic data primarily focused on gender and region. Additional demographic factors could provide a deeper understanding of audience preferences.
2. **External Validity:** The findings are specific to the Slovenian radio broadcasting market and may not make sense for other regions or markets without further investigation.
3. **Client-Specific Requirements:** The research was conducted according to the specifications of our client, which may have influenced the study's direction and the nature of the data collected.
4. **Technological Constraints:** While Power BI provided strong data visualization and analysis capabilities, other analytical tools and methods might produce additional insights.
5. **Survey Methodology:** The reliance on telephone and online surveys may introduce a bias if certain demographics are less likely to participate in such surveys, potentially skewing the results.

5. CONCLUSIONS AND FUTURE WORKS

In this thesis, I explored analyzing the Slovenian radio broadcasting market using Power BI. The focus was on demographic preferences for radio stations. This project's phases, such as data collection, data transformation, visualization, and final analysis, have revealed valuable insights that could greatly influence radio stations' strategies and enhance their ability to adjust their content for specific audiences. The findings that I have highlighted in the 4.1. section of the thesis, which confirm the impact of demographic factors on listening habits, can help radio stations adapt their marketing strategies and advertising approaches.

Through analysis, I discovered that radio station VAL 202 reached the most people both in daily and weekly reach, confirming that it has a strong presence in the market. On the other hand, the reach of Radio Aktual and Radio 1 was only slightly smaller, meaning that the market is dynamic and without a clear winner. This diversity shows that preferences vary among the Slovenian population and opens new opportunities for broadcasters.

Considering that the insights show that some demographics, like older listeners and listeners from lower incomes, are more likely to listen to radio in general, radio stations can adjust their program and marketing to fit with the needs of these demographics. With streaming services becoming more popular and rapidly taking the market share of radio stations, these adaptations are important and necessary for the continuous success of radio broadcasting in general and individual radio stations.

The use of Power BI exhibited its ability to transform raw data into an actionable report. Moreover, it helped the users understand their audience and their preferences more deeply. This project indicates that using analytics in radio broadcasting is beneficial in enhancing strategic decision-making and can act as a basis for further application of business analytics to maintain competitiveness and relevance.

Future Works

The project will go on and more data will be added quarterly to carry out an ongoing analysis of the broadcasting market and help radio stations make better decisions.

Further work could be undertaken using a wider range of demographics, including lifestyle choices, job roles, marital status, and more. This would help gain an even deeper understanding of the audience's preferences. Moreover, in the future, a more time-based analysis and tracking trends over time would be useful and beneficial to help broadcasters understand how preferences are influenced and change over time.

Another interesting area that could be explored is the competition between the radio stations and the streaming services. This could help the radio station understand the business of streaming services better, identify the weak and strong points and help them gain competitive advantage.

Last but not least, more advanced methodologies and analytics could be used to further analyze the data that is already available, to make sure that the complex database is fully utilized.

This suggested future work proposes the expansion of the scope of the project by adding more demographic factors, analyzing the competition with not only other radio stations but also streaming services, and using more advanced analysis for deeper insights. This effort would provide a more detailed perspective on the radio broadcasting industry, empowering stakeholders to anticipate and adapt to market trends effectively.

BIBLIOGRAPHICAL REFERENCES

Agile Broadcast. (2015, December 17). *Types of Radio Broadcasting*. https://www.agilebroadcast.com.au/types-of-radio-broadcasting/

Bailer, W., Wijnants, M., Lievens, H., & Claes, S. (2020). Multimedia Analytics Challenges and Opportunities for Creating Interactive Radio Content. In Y. M. Ro, W.-H. Cheng, J. Kim, W.-T. Chu, P. Cui, J.-W. Choi, M.-C. Hu, & W. De Neve (Eds.), *MultiMedia Modeling* (Vol. 11962, pp. 375–387). Springer International Publishing. https://doi.org/10.1007/978-3-030-37734-2_31

Becker, L. T., & Gould, E. M. (2019). Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Serials Review*, *45*(3), 184–188. https://doi.org/10.1080/00987913.2019.1644891

Britannica. (2024, January 18). *Broadcasting—Radio, TV, Networks | Britannica*. https://www.britannica.com/technology/broadcasting

Hallett, L., & Hintz, A. (2010). Digital broadcasting – Challenges and opportunities for European community radio broadcasters. *Telematics and Informatics*, *27*(2), 151–161. https://doi.org/10.1016/j.tele.2009.06.005

Hirschmeier, S., Tilly, R., & Beule, V. (2019). *Digital Transformation of Radio Broadcasting: An Exploratory Analysis of Challenges and Solutions for New Digital Radio Services*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2019.602

IBM. (2024). *What is ETL (Extract, Transform, Load)? | IBM*. https://www.ibm.com/topics/etl

Ismandianto, I., Suyanto, S., Latifah, K., & Muchid, M. (2022). Transformation of Radio Technology in The Digital Age. *Nyimak: Journal of Communication*, *6*(1), 115. https://doi.org/10.31000/nyimak.v6i1.5547

Krishnan, V., & Chang, S. G. (2000). Customized Internet radio. *Computer Networks*, *33*(1), 609–618. https://doi.org/10.1016/S1389-1286(00)00066-9

Kuyucu, M. (2019, January 1). *Digital Convergence of Radio: Effects of Digitilazation on Radio Media*.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management*, *36*(5), 700–710. https://doi.org/10.1016/j.ijinfomgt.2016.04.013

Marketing Charts. (2016, March 8). How Various Demographic Groups Listen to the Radio. *Marketing Charts*. https://www.marketingcharts.com/demographics-and-audiences/men-demographics-and-audiences-66123

Mehanović, D., & Durmić, N. (2022). Case Study Application of Business Intelligence in Digital Advertising: *International Journal of E-Business Research*, *18*(1), 1–16. https://doi.org/10.4018/IJEBR.293294

Reis, J., & Housley, M. (2022). *Fundamentals of data engineering: Plan and build robust data systems* (First edition). O'Reilly.

Sabharwal, R., & Miah, S. J. (2021). A new theoretical understanding of big data analytics capabilities in organizations: A thematic analysis. *Journal of Big Data*, *8*(1), 159. https://doi.org/10.1186/s40537-021-00543-6

Singh, A. M. (2023, March 27). *What is Radio Advertising? | Definition, Types, and Examples*. https://www.themediaant.com/blog/what-is-radio-advertising/

Sumalatha, V. (2017). *A Review on Different Datamining Schemas with their Design Methodologies*. *19*(5), 5–8.

Techwalla. (n.d.). *Definition of Radio Broadcasting*. Techwalla. Retrieved January 28, 2024, from https://www.techwalla.com/articles/definition-of-radio-broadcasting

WorldDAB. (2024). *DAB | WorldDAB*. https://www.worlddab.org/dab

APPENDIX A

a. How many radio programs have you listened to in the **last seven days**?

1. One
2. Two
3. Three
4. Four
5. Five or more
6. I don't know, I cannot answer
7. None, I never listen to the radio

td. Which (**Slovenian**) radio programs have you listened to in the last week (in the past 7 days) for at least 15 minutes?

Instructions: Please try to recall **all the stations** you have listened to and mark them on the list.

You may indicate up to a maximum of **10 stations.**

1. PRVI (first program RA SLO)
2. VAL 202 (second program RA SLO)
3. ARS (third program RA SLO)
4. Koroški radio
5. Moj radio (Moj radio 1)
6. Muravidek Magyar Radio
7. Primorski val
8. Radio 1
9. Radio 94
10. Radio Aktual
11. Radio Brezje
12. Radio Capodistria
13. Radio Capris
14. Radio Celje
15. Radio CENTER - Hitradio Center
16. Radio City
17. Radio Ekspres
18. Radio Bob
19. Radio Gorenc
20. Radio Hit
21. Radio Koper
22. Radio Kranj
23. Radio Krka

24. Radio Aktual Kum
25. Radio Jam Fm
26. Radio Maribor
27. Radio Marš
28. Radio Maxi (Prleški val)
29. Radio Murski val
30. Radio Net FM
31. Radio Ognjišče
32. Radio Pohorje
33. Radio Prlek (Ormož)
34. Radio Ptuj
35. Radio Robin
36. Radio Rogla
37. Radio Romic
38. Radio Salomon
39. Radio SI (Radio Slovenia International)
40. Radio Slovenske gorice (RSG)
41. Radio Sora
42. Radio Sraka
43. Radio Študent (RŠ)
44. Radio Top
45. Radio Triglav
46. Radio Univox
47. Radio Velenje
48. Radio Veseljak
49. Rock Radio
50. Štajerski val
51. Zeleni val
52. Radio1 80-ta
53. Rock Maribor
54. Radio Antena – Hitradio Antena
55. Radio Fantasi (Antena Štajerska)
56. Rock Celje
57. A Radio
58. Glasbeni radio
59. Best FM
60. I don't know.
61. Other (1st mention):
62. Other (2nd mention):
63. Other (3rd mention):
64. I did not listen to any radio station last week.

j. Did you listen to the radio for at least fifteen minutes yesterday?

Try to remember and respond, even if you listened to the radio anywhere and it could also be just in passing, for example, at work, school, on the bus, at a friend's, visiting, in a bar, etc.

1. Yes
2. No

k. Are you certain that you DID NOT listen to the radio even for fifteen minutes yesterday?

1. I did NOT listen to the radio yesterday.
2. I DID listen to the radio for at least 15 minutes yesterday.

If the respondent did listen to the radio yesterday:

You've mentioned that you listened to the radio yesterday. Please recall YESTERDAY and specify which radio programs you listened to and the exact time frame of listening.

Detail Request: Since we are interested in radio program listening in 15-minute increments, please define your listening experience as precisely as possible.

_____

ez. Your gender:

1. Male
2. Female

st. Can you share your year of birth with us? (For example, 1978) _____

fb. What is your highest completed level of education?

1. Incomplete primary education
2. Completed primary education
3. Completed 2 or 3-year vocational school
4. Completed 4-year high school
5. Completed 2-year higher education, 3-year college education (Bologna process 6th level)
6. Completed university degree, Bologna master's degree or higher (Bologna process 7th level or higher)

7. Prefer not to answer

dohodekosebni. The average net monthly income in Slovenia is around 1,320 EUR. Would you say your net monthly income is below average, average, or above average?

1. Below average income (less than 1320€)
2. Average income (around 1320€)
3. Above average income (more than 1320€)
4. I do not have a personal income
5. Prefer not to say

tipnaselja. What type of settlement do you live in?

1. Urban (city environment)
2. Rural (village environment)
3. I don't know, I prefer not to say

regija. Please indicate the statistical region in which you reside.

1. Pomurska (Mura)
2. Podravska (Drava)
3. Koroška (Carinthia)
4. Savinjska (Savinja)
5. Zasavska (Central Sava)
6. Posavska (Lower Sava)
7. Jugovzhodna Slovenija (Southeast Slovenia)
8. Osrednjeslovenska (Central Slovenia)
9. Gorenjska (Upper Carniola)
10. Primorsko-notranjska (Littoral–Inner Carniola)
11. Goriška (Gorizia)
12. Obalno-kraška (Coastal–Karst)

obcina. Municipality in which you reside: _____

APPENDIX B

Fact Tables

fact PTE Poslusanost tedenski

Column Name	Data Type
PTE ID baze	Whole Number
PTE Mesec	Whole Number
PTE Leto	Whole Number
PTE Dan anketiranja	Whole Number
PTE Anketiranec ID	Whole Number
PTE Spol	Whole Number
PTE Leto rojstva	Whole Number
PTE Izobrazba	Whole Number
PTE Dohodkovni razred	Whole Number
PTE Tip naselja	Whole Number
PTE Regija	Whole Number
PTE Obcina	Whole Number
PTE Radio ID	Whole Number
PTE Utez	Decimal Number

fact PDN Poslusanost dnevni

Column Name	Data Type
PDN Leto	Whole Number
PDN mesec	Whole Number
PDN Dan anketiranja	Whole Number
PDN Anketiranec ID	Whole Number
PDN Spol	Whole Number
PDN Leto rojstva	Whole Number
PDN Izobrazba	Whole Number
PDN Dohodkovni razred	Whole Number
PDN Tip naselja	Whole Number
PDN Regija	Whole Number
PDN Obcina	Whole Number
PDN Casovni interval	Text
PDN Radio ID	Whole Number
PDN Utez	Decimal Number
PDN ID baze	Whole Number

fact PRE Prebivalstvo

Column Name	Data Type
PRE Občina	Text
PRE Starost	Whole Number
PRE Spol	Text
PRE Izobrazba	Text
PRE Dohodkovni razred	Text
PRE Vrednost	Whole Number

fact PIN Poslusanost intervali

Column Name	Data Type
PIN ID baze	Whole Number
PIN Mesec	Whole Number
PIN Leto	Whole Number
PIN Dan anketiranja	Whole Number
PIN Anketiranec ID	Whole Number
PIN Spol	Whole Number
PIN Leto rojstva	Whole Number
PIN Izobrazba	Whole Number
PIN Dohodkovni razred	Whole Number
PIN Tip naselja	Whole Number
PIN Regija	Whole Number
PIN Obcina	Whole Number
PIN Casovni interval	Text
PIN Radio ID	Whole Number
PIN Utez	Decimal Number

Dimension Tables

dim INT Casovni intervali

Column Name	Data Type
INT ID	Text
INT Interval	Text
INT Sort	Whole number
INT Interval ura	Text
INT Interval dve uri	Text
INT Sort ura	Whole number
INT Sort dve uri	Whole number

dim DOH Mesecni dohodek

Column Name	Data Type
DOH ID	Whole Number
DOH Mesecni dohodek	Text

dim TNA Tip naselja

Column Name	Data Type
TNA ID	Whole Number
TNA Tip naselja	Text

dim IZB Izobrazba

Column Name	Data Type
IZB ID	Whole Number
IZB Izobrazba	Text

dim SPO Spol

Column Name	Data Type
SPO ID	Whole Number
SPO Spot	Text

dim RAD Radio

Column Name	Data Type
TNA ID	Whole Number
TNA Tip naselja	Text

dim OBC Obcine

Column Name	Data Type
OBC Obcina ID	Whole number
OBC Naziv	Text
OBC Regija ID	Whole number
OBC Regija naziv	Text

APPENDIX C

Calculated Tables

dim DAA Dan anketiranja

Column Name	Data Type
DAA Dan anketiranja	Whole Number
DAA Dan poslušanja	Whole Number
DAA Dan	Text

dim IDB ID baze

Column Name	Data Type
IDB ID Baze	Whole Number

dim STR Starost

Column Name	Data Type
STR Starost	Whole Number
STR Starostna skupina	Text
STR Starost posodobljeno	Whole Number

dim KOL Koledar

Column Name	Data Type
KOL Datum	Date
KOL Mesec ID	Whole Number
KOL Mesec	Text
KOL Leto	Whole Number
KOL Kvartal	Text
KOL Leto-mesec	Text
KOL Leto-kvartal	Text
KOL Leto-mesec ID	Date

APPENDIX D

CALCULATED TABLES

dim DAA Dan anketiranja

The table was created to establish a link between *fact PTE Poslusanost tedenski*, *fact PIN Poslusanost intervali*, and *fact PDN Poslusanost dnevni*. It was generated using the following formula:

```
dim DAA Dan anketiranja =
DISTINCT ( 'fact PTE Poslusanost tedenski'[PTE Dan anketiranja] )
```

This DAX formula created the table from all days in the week on which surveys were conducted from the *fact PTE Poslusanost tedenski* table. This was to create a dimension that can be used to filter or group fact data in the report.

dim IDB ID baze

Much like the table above, this table was also created to provide a connection between *fact PTE Poslusanost tedenski*, *fact PIN Poslusanost intervali*, and *fact PDN Poslusanost dnevni*. The DAX formula for this table is the following:

```
dim                    IDB                    ID                    baze                    =
DISTINCT( 'fact PTE Poslusanost tedenski'[PTE ID baze] )
```

Similarly to the previous DAX formula, this formula generated a table of unique IDs of databases from the *fact PTE Poslusanost tedenski* table. The purpose of the table was also to help filter the data from the fact tables.

dim KOL Koledar

Since a cruicial part of our analysis was following data based on time, it was necessary to create a calendar table that could connect to the fact tables which included a date column. These tables were again *fact PTE Poslusanost tedenski*, *fact PIN Poslusanost intervali*, and *fact PDN Poslusanost dnevni*. The DAX used to create the table was the following:

```
dim KOL Koledar =
CALENDAR (
    DATE ( MIN ( 'fact PTE Poslusanost tedenski'[PTE Leto] ), 1, 1 ),
    DATE ( MAX ( 'fact PTE Poslusanost tedenski'[PTE Leto] ), 12, 31 )
)
```

The formula constructed a date table that spanned from January 1st of the earliest year found in the *[PTE Leto]* column of the *fact PTE Poslusanost tedenski* table to December 31st of the latest year found in the same column.

From here, we added more columns to have more options of displaying the data. The columns added were *[KOL Mesec ID], [KOL Mesec], [KOL Leto], [KOL Kvartal], [KOL Leto-mesec], [KOL Leto-kvartal],* and *[KOL Leto-mesec ID].* The columns were added using the following formulas:

```
KOL Mesec ID =
MONTH ( 'dim KOL Koledar'[KOL Datum] )

KOL Mesec =
SWITCH (
    [KOL Mesec ID],
    1, "Januar",
    2, "Februar",
    3, "Marec",
    4, "April",
    5, "Maj",
    6, "Junij",
    7, "Julij",
    8, "Avgust",
    9, "September",
    10, "Oktober",
    11, "November",
    12, "December",
    BLANK ()
)

KOL Leto =
YEAR ( 'dim KOL Koledar'[KOL Datum] )

KOL Kvartal =
"Q" & QUARTER ( 'dim KOL Koledar'[KOL Datum] )

KOL Leto-mesec =
'dim KOL Koledar'[KOL Leto] & " - " & 'dim KOL Koledar'[KOL Mesec]

KOL Leto-kvartal =
'dim KOL Koledar'[KOL Leto] & " - " & 'dim KOL Koledar'[KOL Kvartal]

KOL Leto-mesec ID =
FORMAT ( 'dim KOL Koledar'[KOL Datum], "YYYY - MM" )
```

dim STR Starost

A final table was added to categorize survey respondents by age groups. This provided a structured approach to analyze listenership trends across different age demographics. It was generated with the following DAX formula:

```
dim STR Starost =
GENERATESERIES (
    MIN ( 'fact PTE Poslusanost tedenski'[PRA Starost] ),
    MAX ( 'fact PTE Poslusanost tedenski'[PRA Starost] ),
    1
```

```
)
```

The formula resulted in a calculated table that listed individual ages in ascending order, beginning from the youngest to the oldest age recorded in the *[PRA Starost]* column of the *fact PTE Poslusanost tedenski* table, with the sequence advancing by one year at a time.

An additional column was added to categorize the ages into age groups.

```
STR Starost posodobljeno =
IF (
    'dim STR Starost'[STR Starost] < 15,
    15,
    IF ( 'dim STR Starost'[STR Starost] > 80, 80, 'dim STR Starost'[STR Starost] )
)
```

CALCULATED COLUMNS

fact PIN Poslusanost intervali

Given that the table lacked a date column and therefore couldn't be linked to our calendar table, we added a new calculated column called [PIN Datum obdobja]. The column was added with the DAX formula below:

```
PIN Datum obdobja =
DATE ( 'fact PIN Poslusanost intervali'[PIN Leto], 'fact PIN Poslusanost intervali'[PIN
Mesec], 1 )
```

The formula calculates a date by combining the year and month from the *fact PIN Poslusanost intervali* table, setting the day to the first of the month.

fact PDN Poslusanost dnevni

Like the previous table, this one also lacked a date column. Therefore, we had to add it by calculating the column ourselves. The DAX formula followed the same approach as in the previous table:

```
PDN Datum obdobja =
DATE ( 'fact PDN Poslusanost dnevni'[PDN Leto], 'fact PDN Poslusanost dnevni'[PDN Mesec], 1
)
```

Another calculated column was added to calculate the age of the respondent, so it could be linked with the *dim STR Starost* table. The DAX is the following:

```
PDN Starost =
IF (
    'fact PDN Poslusanost dnevni'[PDN Leto rojstva] = BLANK (),
    BLANK (),
    'fact PDN Poslusanost dnevni'[PDN Leto] - 'fact PDN Poslusanost dnevni'[PDN Leto rojstva]
)
```

The formula calculates the age by subtracting the birth year [PDN Leto rojstva] from the current year [PDN Leto] in the *fact PDN Poslusanost dnevni* table. It returns a blank value if the birth year is not specified, ensuring data integrity by only calculating the age where birth year data is available.

fact PTE Poslusanost tedenski

Further transformations applied to this table mirrored those performed on the *PDN Poslusanost dnevni* table, using the same two DAX formulas for calculation:

```
PTE Datum obdobja =
DATE ( 'fact PTE Poslusanost tedenski'[PTE Leto], 'fact PTE Poslusanost tedenski'[PTE Mesec],
1 )

PTE Starost =
IF (
    'fact PTE Poslusanost tedenski'[PTE Leto rojstva] = BLANK (),
    BLANK (),
     'fact PTE Poslusanost tedenski'[PTE Leto] - 'fact PTE Poslusanost tedenski'[PTE Leto
rojstva]
)
```

dim IZB Izobrazba

To this table, we added a calculated column to categorize education levels into different groups. The following DAX formula was used:

```
IZB Izobrazba skupine =
SWITCH (
    'dim IZB Izobrazba'[IZB ID],
    1, "Osnovnošolska ali manj",
    2, "Osnovnošolska ali manj",
    3, "Srednješolska",
    4, "Srednješolska",
    5, "Višješolska, visokošolska ali več",
    6, "Višješolska, visokošolska ali več",
    7, "Ne želim odgovoriti"
)
```

This DAX formula classifies individual education level into specific groups using the [IZB ID] column. The `SWITCH` function simplifies the process by linking each education level ID to its designated group, effectively organizing education levels into categories.

CALCULATED MEASURES

fact PDN Poslusanost dnevni

```
PDN Doseg % =
( DIVIDE ( 'fact PDN Poslusanost dnevni'[PDN Doseg N], [PTE Numerus skupaj] ) )

PDN Doseg % povzetek =
DIVIDE (
    [PDN Doseg %],
    CALCULATE ( [PDN Doseg %], ALL ())
)

PDN Doseg % skupaj =
DIVIDE ( [PDN Doseg N skupaj], [PTE Numerus skupaj] )

PDN Doseg N =
IF (
    [PTE Numerus skupaj] < 60,
    BLANK (),
    CALCULATE (
        CALCULATE (
            SUM ( 'fact PDN Poslusanost dnevni'[PDN Utez] ),
            NOT ( 'dim RAD Radio'[RAD Radio ID] IN { 99, 96, 97, 98, 998 } )
        )
    )
)

PDN Doseg N skupaj =
        IF(
            [PTE Numerus skupaj]<60,
            BLANK(),
            CALCULATE(
               SUMX(DISTINCT('fact PDN Poslusanost dnevni'[PDN Anketiranec ID]),[PDN Maxx]),
                NOT('fact PDN Poslusanost dnevni'[PDN Radio ID] in {99,96,97,98,998}),
                'dim IDB ID baze'[IDB ID baze]>MAX('dim IDB ID baze'[IDB ID baze])-4

            )
    )

PDN DT Naslov =
"Dnevna poslušanost za "
    & FIRSTNONBLANK ( 'dim RAD Radio'[RAD Naziv], BLANK () = FALSE () )
PDN Maxx =
MAXX (
    DISTINCT ( 'fact PDN Poslusanost dnevni'[PDN Anketiranec ID] ),
    MAX ( 'fact PDN Poslusanost dnevni'[PDN Utez] )
)
```

fact PIN Poslusanost intervali

```
PIN Average Doseg N =
AVERAGEX ( 'dim INT Casovni intervali', [PIN Doseg N] )

PIN Doseg % =
IF (
    ISFILTERED ( 'dim DAA Dan anketiranja'[DAA Dan] ),
    DIVIDE (
        'fact PIN Poslusanost intervali'[PIN Average Doseg N],
        [PIN Numerus skupaj]
    ),
    DIVIDE (
        'fact PIN Poslusanost intervali'[PIN Average Doseg N],
        [PTE Numerus skupaj]
    )
)

PIN Doseg % 1-5 =
CALCULATE (
    [PIN Doseg %],
    NOT ( 'dim DAA Dan anketiranja'[DAA Dan poslušanja] IN { 6, 7 } )
)

PIN Doseg % 1-7 =
DIVIDE (
    'fact PIN Poslusanost intervali'[PIN Average Doseg N],
    [PTE Numerus skupaj]
)

PIN Doseg % 6-7 =
CALCULATE (
    [PIN Doseg %],
    NOT ( 'dim DAA Dan anketiranja'[DAA Dan poslušanja] IN { 1, 2, 3, 4, 5 } )
)

PIN Doseg N =
IF (
    [PTE Numerus skupaj] < 60,
    BLANK (),
    CALCULATE (
        AVERAGEX (
            'dim INT Casovni intervali',
            SUM ( 'fact PIN Poslusanost intervali'[PIN Utez] )
        ),
        NOT ( 'dim RAD Radio'[RAD Radio ID] IN { 99, 96, 97, 98, 998 } ),
        'fact PIN Poslusanost intervali'[PIN ID baze]
            > MAX ( 'fact PIN Poslusanost intervali'[PIN ID baze] ) - 4
    )
)
```

```
PIN Numerus skupaj =
DIVIDE (
    CALCULATE (
        DISTINCTCOUNT ( 'fact PTE Poslusanost tedenski'[PTE Anketiranec ID] ),
        ALLSELECTED (),
        ALL ( 'dim RAD Radio' ),
        'dim IDB ID baze'[IDB ID baze]
            > MAX ( 'dim IDB ID baze'[IDB ID baze] ) - 4
    ),
    7
)
```

fact PRE Prebivalstvo

```
PRE Doseg dnevni v 1000 =
ROUND (
     'fact PRE Prebivalstvo'[PRE Število prebivalcev] * 'fact PDN Poslusanost dnevni'[PDN
Doseg %] / 1000,
    0
)

PRE Doseg dnevni v 1000 skupaj =
ROUND (
    'fact PRE Prebivalstvo'[PRE Število prebivalcev] * [PDN Doseg % skupaj] / 1000,
    0
)

PRE Doseg tedenski v 1000 =
ROUND (
    'fact PRE Prebivalstvo'[PRE Število prebivalcev] * 'fact PTE Poslusanost tedenski'[PTE
Doseg %] / 1000,
    0
)

PRE Število prebivalcev =
CALCULATE (
    SUM ( 'fact PRE Prebivalstvo'[PRE Vrednost] ),
    ALLSELECTED ( 'dim RAD Radio'[RAD Naziv] ),
    ALLSELECTED ( 'dim OBC Obcine'[OBC Regija naziv] ),
    ALLSELECTED ( 'dim DOH Mesecni dohodek'[DOH Mesecni dohodek] ),
    ALLSELECTED ( 'dim STR Starost'[STR Starostna skupina] ),
    ALLSELECTED ( 'dim SPO Spol'[SPO Spol] ),
    ALLSELECTED ( 'dim IZB Izobrazba'[IZB Izobrazba skupine] ),
    ALLSELECTED ( 'dim STR Starost'[STR Starost posodobljeno] ),
    ALLSELECTED ( 'dim OBC Obcine'[OBC Obcina ID] )
)
```

fact PTE Poslusanost tedenski

```
PTE barva =
IF ( [PTE Numerus] < 60, "#E00040", "#FFFFFF" )
```

```
PTE Doseg % =
ROUNDUP (
    DIVIDE ( 'fact PTE Poslusanost tedenski'[PTE Doseg N], [PTE Numerus skupaj] ),
    3
)

PTE Doseg N =
CALCULATE (
    SUM ( 'fact PTE Poslusanost tedenski'[PTE Utez] ),
    NOT ( 'dim RAD Radio'[RAD Radio ID] IN { 99, 96, 97, 98, 998 } )
)

PTE Doseg N skupaj =
CALCULATE (
    COUNT ( 'fact PTE Poslusanost tedenski'[PTE ID baze] ),
    ALL ( 'dim RAD Radio' ),
    NOT ( 'dim RAD Radio'[RAD Radio ID] IN { 99, 96, 97, 98, 998 } )
)

PTE Numerus =
CALCULATE (
    DISTINCTCOUNT ( 'fact PTE Poslusanost tedenski'[PTE Anketiranec ID] )
)

PTE Numerus naslov =
IF (
    [PTE Numerus] > 60,
    "",
    "Premajhen numerus, statistično sklepanje ni mogoče."
)

PTE Numerus skupaj =
CALCULATE (
    DISTINCTCOUNT ( 'fact PTE Poslusanost tedenski'[PTE Anketiranec ID] ),
    ALLSELECTED (),
    ALL ( 'dim RAD Radio' )
)

PTE Obdobje baz naslov =
IF (
    [PTE Numerus] = BLANK (),
    BLANK (),
    "Od " & FIRSTNONBLANK ( 'dim KOL Koledar'[KOL Leto-mesec ID], [PTE Numerus] ) & " do "
        & LASTNONBLANK ( 'dim KOL Koledar'[KOL Leto-mesec ID], [PTE Numerus] )
)
```

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa