

PROFESSIONAL TRAINING REPORT
Sathyabama Institute of Science and Technology
(Deemed to be University)

Submitted in partial fulfilment of the requirements for the award
of Bachelor of Technology Degree in Information Technology

By

Dhanush Raj M

(40120027)



DEPARTMENT OF INFORMATION TECHNOLOGY
SCHOOL OF COMPUTING
SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

JEPPIAAR NAGAR, RAJIV GANDHI SALAI,
CHENNAI, TAMILNADU – 600119

April 2023



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

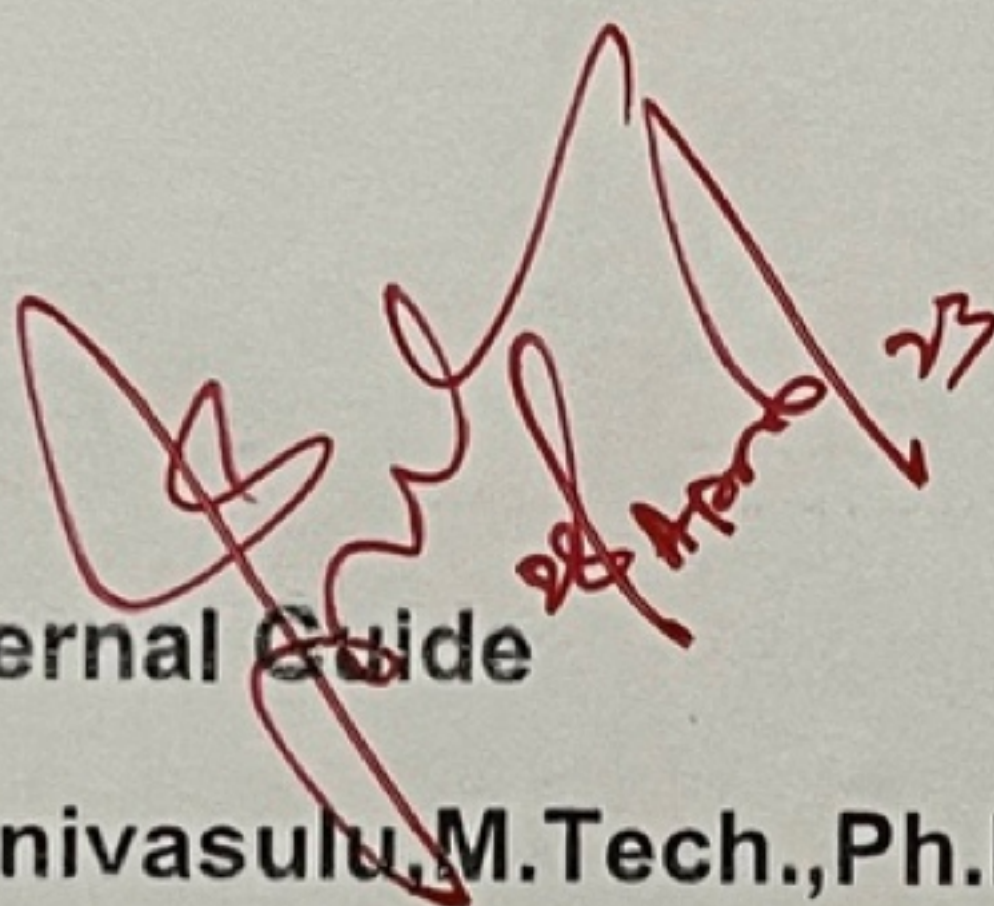
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF INFORMATION TECHNOLOGY

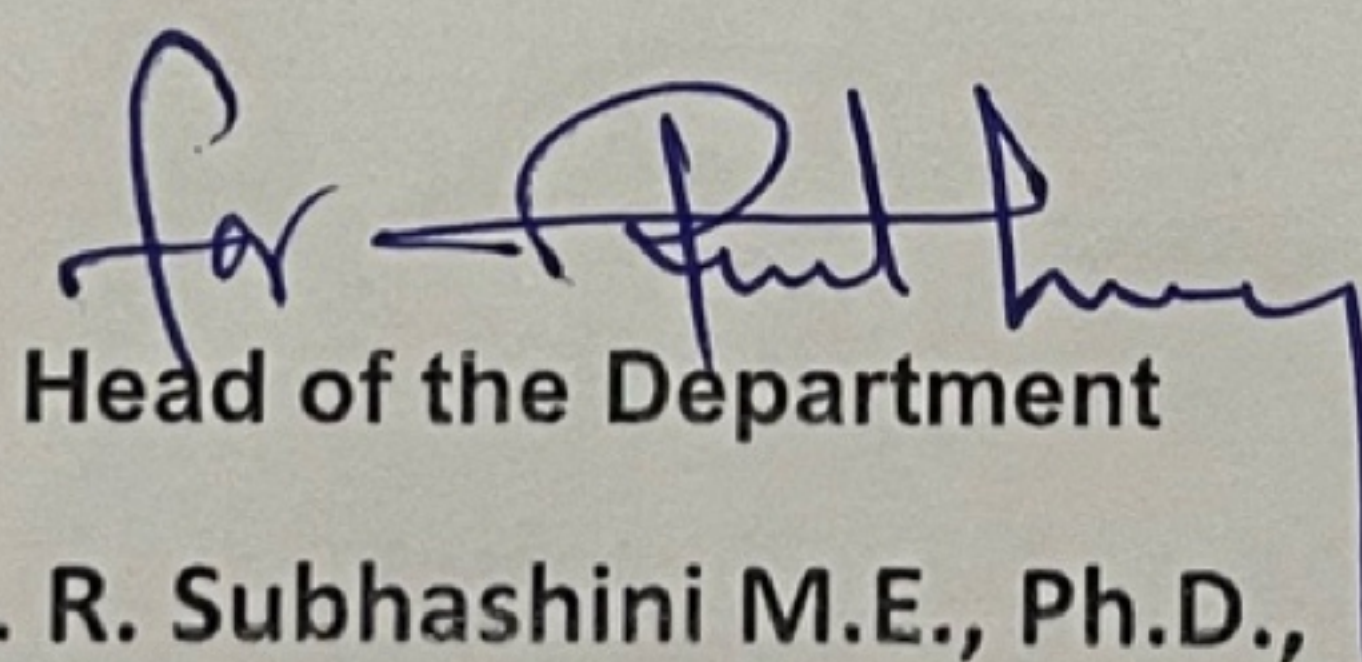
BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **DHANUSH RAJ M** who carried out the project entitled "**MEDICAL INSURANCE CLAIMS PREDICTION**" under my supervision from August 2022 to October 2022.



Internal Guide

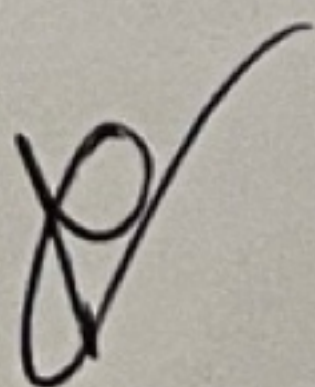
Dr. Senduru Srinivasulu, M.Tech., Ph.D.,



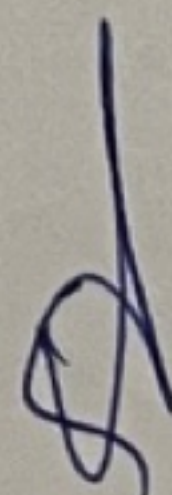
Head of the Department

Dr. R. Subhashini M.E., Ph.D.,

Submitted for Viva Voce Examination held on 28/04/2023



Internal Examiner



External Examiner

DECLARATION

I **DHANUSH RAJ M** hereby declare that the Project Report entitled "**MEDICAL INSURANCE CLAIMS PREDICTION**" done by me under the guidance of **Dr.Senduru Srinivasulu M.Tech.,Ph.D.**, is submitted in partial fulfillment of the requirements for the award of Bachelor of Technology degree in Information Technology.

DATE: 28/04/2023

M. Dhah

PLACE: CHENNAI

SIGNATURE OF CANDIDATE:

ACKNOWLEDGEMENT

I am pleased to acknowledge our sincere thanks to Board of management of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey our thanks to **Dr. T. SASIKALA M.E., Ph.D., Dean, School of Computing** and **Dr. R. SUBHASHINI M.E., Ph.D., Head of the Department, Department of Information Technology** for providing us the necessary support and details at the right time during the progressive reviews.

I would like to express our sincere and deep sense of gratitude to our Project, **Dr.Senduru Srinivasulu M.Tech.,Ph.D.**, for her valuable guidance, suggestions and constant encouragement paved the way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Information Technology** who were helpful in many ways for the completion of the project.

CERTIFICATE :

No. : 34337

 **राष्ट्रीय लघु उद्योग निगम लिमिटेड**
NATIONAL SMALL INDUSTRIES CORPORATION LIMITED
(A Govt. of India Enterprise)

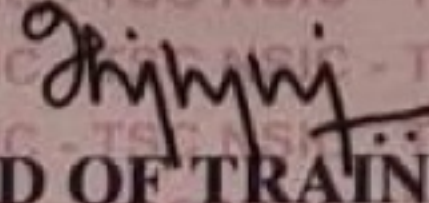
एन एस आई सी
N S I C

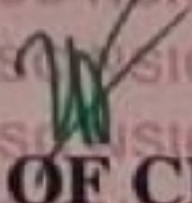
NSIC - TECHNICAL SERVICES CENTRE
Sector B-24, Guindy Industrial Estate, Ekkaduthangal, Chennai - 600 032.

CERTIFICATE

This is to certify that Mr. **DHANUSH RAJ M.** S/o.
Shri. **MADHUSUDAN S.** student of III year B.Tech.
(IT), Sathyabama Institute of Science and
Technology has undergone "**SKILL DEVELOPMENT
PROGRAM ON DATA ANALYTICS**" conducted by
us for a period of one week (30 Hours) from
09.03.2023 to 14.03.2023.




HEAD OF TRAINING


HEAD OF CENTRE

ABSTRACT

Health claims contain a wealth of information that can be turned into valuable insights for insurers. The potential of analytics in this domain goes beyond claims management, helping the industry reduce fraud and make healthcare payments more transparent. In addition, almost 70% of life insurers admit that predictive analysis helps them reduce expenses and improve sales. But despite this fact, data analytics in claims management is still underestimated. Information derived from health claims and insurance plans data can provide insights into a patient and a population health status and inform healthcare plans and providers on patient health trajectories, potential interventions and cost efficiency. In this project, we have taken a medical insurance claims dataset and analyzed it to get insights using Microsoft Power BI which is a data visualization software which also helps in prediction using the data provided.

TABLE OF CONTENTS

	ABSTRACT	v
	LIST OF FIGURES	vii
1	INTRODUCTION	1
	1.1 FORECASTING	2
	1.2 CLAIMS FORECASTING	2
	1.3 CLAIMS PREDICTION	3
	1.4 DATA PROCESSING	4
	1.5 SPECIFICATION	5
2	LITERATURE SURVEY	5
3	METHODS AND ALGORITHM USED	10
	3.1 PROJECT DEVELOPMENT	10
	3.2 DATA ANALYTICS	10
	3.3 DESCRIPTION ANALYSIS	11
	3.4 SYSTEM REQUIREMENTS	16
	3.5 POWER BI	17
	3.6 METHODOLOGY	18
4	RESULTS AND DISCUSSION	27

5	CONCLUSION	30
	REFERENCES	31
	APPENDIX	32

LIST OF FIGURES

FIGURE NO	FIGURE NAME	PAGE NO
3.1	PROJECT DEVELOPMENT	10
3.2	EXCEL DATA	22
3.3	CLAIMED AMOUNT BY PROCEDURE WISE	23
3.4	CLAIMED AMOUNT BY BRANCH WISE	24
3.5	AMOUNT CLAIMED BY TIME TAKEN	25
3.6	CLAIMS APPROVED/DENIED	25
3.7	INSURANCE CLAIMED BY AGE	26
4.1	MEDICAL INSURANCE CLAIMS IN POWER BI	27

CHAPTER-1

INTRODUCTION

Medical insurance claim is a formal request made by health care providers or policyholders to their insurance companies for reimbursement of medical expenses. It is a critical process in the health care industry that helps patients receive the medical care they need while ensuring health care providers are appropriately compensated for their services.

The medical insurance claim process involves detailed documentation and communication between health care providers, patients, and insurance companies. It typically includes information such as the diagnosis, medical treatment provided, and the cost of services. Once the claim is submitted, the insurance company will review the claim to determine if the medical treatment is covered under the policy, and if so, how much the insurance company will pay.

Medical insurance claims can be complex, and the process may vary depending on the type of insurance policy and the medical treatment provided. It is important for health care providers to understand the claims process and to ensure that they submit accurate and complete information to insurance companies to avoid claim denials or delays in payment.

The process of submitting a medical insurance claim involves providing detailed information about the medical services provided, including the diagnosis, treatment codes, and associated costs. The insurance company then reviews the claim to determine if the services provided are covered under the patient's insurance policy and if the provider's charges are reasonable.

Overall, medical insurance claims play a critical role in the health care industry by enabling patients to receive the medical care they need while also ensuring that health care providers are fairly compensated for their services.

1.1 FORECASTING

Forecasting is the process of projecting the claimed amounts. Implementing a forecasting system enables you to assess current market trends and claims quickly so that you can make informed decisions about the operations. It is a planning tool that enables businesses to chart their next moves and create budgets that will hopefully cover whatever uncertainties may occur.

1.1.1 Types of Forecasting

- Quantitative
- Qualitative

1.2 CLAIMS FORECASTING

Claims forecasting is a prediction of future claims revenue and claims forecasts are usually based on historical data, industry trends, and the status of the current claims pipeline. Businesses use the claims forecast to estimate weekly, monthly, quarterly, and annual claims totals. Just like a weather forecast, your team should view your claims forecast as a plan to work from, not a firm prediction. Claims forecasting is also different from claims goal-setting. While a claims goal describes what you want to happen, a claims forecast estimates what will happen, regardless of your goal. They can also help enterprises to plan their marketing and advertising efforts. Claims forecasting can help businesses make informed decisions about allocating resources. Claims forecasting is not an exact science, and uncertainty is always involved. However, claims forecasts can give businesses a good idea of what to expect in the future and help them plan accordingly.

1.2.1 Pros and Cons

Pros:

- i) Provides insights into upcoming claims and revenue.
- ii) Appeals to stakeholders.
- iii) Aligns business strategies and result.
- iv) Provides opportunities to make adjustments.

- v) Reveals patterns in data.
- vi) Helps integrate lead sources.
- vii) Begins gathering data on new products or services.
- viii) Establishes a claims process.

Cons:

- i) Is hard to predict.
- ii) Creates biased opinions.
- iii) Requires clean data.
- iv) Requires time.
- v) Lacks essential details with numerical data.
- vi) Affected by employee turnover.
- vii) Promotes false optimism.

1.3 CLAIMS PREDICTION

Claims prediction can refer to the process of using data and algorithms to forecast or estimate the likelihood of an event occurring based on historical patterns or other relevant factors.

In the context of insurance, for example, claims prediction involves using data on past claims, policyholders' characteristics, and external factors such as weather and economic conditions to predict the likelihood of future claims. This can help insurance companies better manage risk and make more informed decisions about pricing and coverage. Claims prediction may be used to forecast the likelihood of fraudulent activity, equipment failures, or other events that could have a negative impact on an organization's operations. By identifying and mitigating potential risks before they occur, claims prediction can help organizations save time, money, and resources.

1.3.1 BENEFITS OF CLAIMS PREDICTION:

- Set better goals
- Set case reserves more accurately
- Plans effectively
- Makes decision with confidence

1.4 DATA PROCESSING:

The influencing factors of sales from five aspects: product attributes, merchant attributes, buyer attributes.

1.4 SPECIFICATIONS

1.4.1 HARDWARE CONFIGURATION

- Processor - I5
- Speed - 3 GHz
- RAM - 8GB (min)
- Hard Disk - 500 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - LCD
- Ethernet connection :(LAN)

1.4.2 SOFTWARE CONFIGURATION

- Microsoft Excel sheet
- Power BI Application
- Web Browser

CHAPTER - 2

LITERATURE SURVEY

2.1 Challenges of Using Medical Insurance Claims Data for Utilization Analysis

Research use of insurance claims data presents unique challenges and requires a series of value judgments that are intended to improve the data quality. In this study, medical insurance claims from 2 large companies were combined to assess utilization of complementary and alternative medicine. Challenges included assessing and improving the quality of data, combining data from 2 different companies with dissimilar coding systems, and determining the most appropriate ways to describe utilization. This article addresses 4 methodologic challenges in creating the analytic files: (1) conversion of claims into unique visits, (2) identification of incomplete claims data, (3) categorization of providers and locations of service, and (4) selecting the most useful measures of utilization and expenditures.

2.2 Medical Research Using National Health Insurance Claims Database

All Korean citizens should join the National Health Security System by law. The National Health Insurance Service (NHIS) and the Health Insurance Review and Assessment Service (HIRA) are one of major components to support this system, and all data about medical expenses for the medical claims are stored and managed in the institutions. Recently, medical research using administrative claims databases has dramatically progressed in Korea and worldwide, and the methods how to use them are briefly reviewed in this article. Research using these databases have several strengths. Researchers can perform the complete enumeration survey in a real world. They can get new valuable findings because the number in the database is usually large enough to detect the minute difference with a big statistical power. They can obtain more detailed and reproducible results. Moreover, they can investigate a very rare disease or infrequent side effects of drugs. However, we must recognize that research using administrative claims database also has several

incoherent limitations. These databases have not been constructed originally for research, but for reimbursement. Therefore, there are no important data including medical history and laboratory findings of each patient, which are crucial to adjust baseline characteristics. In addition, it is hard to discover causal relationship and direct association with the included information. In spite of limitations, researchers can easily use these databases for their research now than ever, and the results may be utilized not only to expand the academic knowledge but also to influence the determination of national health care policy. adapting these complex models. Therefore, one interesting aspect is to explore models which are accurate and at the same time applicable for the daily work of fashion companies. Furthermore, it is worth to focus more on the diverse impacting factors which have been described. Different aspects such as the general handling and the reduction possibilities on the forecast results can be examined. Finally, the measurement of the predictive value of user-generated content in real life is a valuable aspect, which is currently followed and explored by the authors of this review.

2.3 Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance

Technological advances in big data (large amounts of highly varied data from many different sources that may be processed rapidly), data sciences and artificial intelligence can improve health-system functions and promote personalized care and public good. However, these technologies will not replace the fundamental components of the health system, such as ethical leadership and governance, or avoid the need for a robust ethical and regulatory environment. In this paper, we discuss what a robust ethical and regulatory environment might look like for big data analytics in health insurance, and describe examples of safeguards and participatory mechanisms that should be established. First, a clear and effective data governance framework is critical. Legal standards need to be enacted and insurers should be encouraged and given incentives to adopt a human-centred approach in the design and use of big data analytics and artificial intelligence. Second, a clear and accountable process is necessary to explain

what information can be used and how it can be used. Third, people whose data may be used should be empowered through their active involvement in determining how their personal data may be managed and governed. Fourth, insurers and governance bodies, including regulators and policy-makers, need to work together to ensure that the big data analytics based on artificial intelligence that are developed are transparent and accurate. Unless an enabling ethical environment is in place, the use of such analytics will likely contribute to the proliferation of unconnected data systems, worsen existing inequalities, and erode trustworthiness and trust. based on its characteristics. Thereafter, once normal predictions are obtained, a regression approach is proposed to model the crisis' impact on sales and correct the predictions accordingly. The proposed correction model is optimized within the time window made up of the few weeks following the start of the crisis, where real sales (affected by the pandemic) and previously predicted sales are available. By analyzing sales behaviors in this window, and optimizing the regression models' parameters, the proposed model is able to predict an optimal correcting coefficient capable of correcting the predicted sales for the next weeks (short-term prediction). We tested the proposed methodology on four different datasets of online sales corresponding to four different families of products. The results showed that we were able to correct sales predictions on the desired period (four weeks) in an efficient way for all of the datasets. Hence, our proposed framework has offered an effective tool to our fashion retailer partner to better manage his supply chain, by knowing which items from the previous collection must be kept in stores, and which new items must be supplied to stores after the end of the lockdown period.

2.4 An emerging trend of big data analytics with health insurance in India

Hugh the elder Chamberlen from Peter Chamberlen family planned construct of insurance within the year 1694. The "Accident Assurance" which began to be offered in nineteenth Century has operated very similar to modern social insurance. This model of payment continued till the beginning or starting of twentieth century. The fashionable insurance programmers evolved from ancient insurance in and throughout the centre to late twentieth

century. Today, the price routine has been cowed by most comprehensive insurance programmers as well as preventive and emergency health care procedures and conjointly most pharmaceuticals. However this is often not forever the case. India is among robust countries with billion and folks, one of the world's quickest growing economy, 29.5% of population are below poverty level, forty sixth of kids square measure underfed yet. The health care structure within the country has been overburdened by the ever increasing population in India. Health care in our country is government supported and government run. But, for many people living within the remote components of the country, accessing primary health care remains a challenge. For the developing countries like India, the health care is one among the foremost important areas to stress, therefore on offer higher health care access to the valuable human resources, that successively can create the India healthier

2.5 Key considerations when using health insurance claims data in advanced data analyses

Online business has grown exponentially during the last decade, and the industries are focusing on online business more than before. However, just setting up an online store and starting selling might not work. Different machine learning and data mining techniques are needed to know the users' preferences and know what would be best for business. According to the decision-making needs of online product sales, combined with the influencing factors of online product sales in various industries and the advantages of deep learning algorithm, this paper Health claims have become a popular source of data for health care analytics, with numerous applications ranging from disease burden estimation and policy evaluation to drug event detection and advanced predictive analytics. Independent of the application, a researcher utilising claims information will likely encounter challenges in using the data, which include dealing with several coding systems and coding irregularities.

Claims-based studies can be much more inclusive than medical records, and they are significantly less expensive. As claims data contain rich cost information concerning medical services, researchers have used such data widely for studies of health care utilization. We highlight some of these challenges and approaches for successful analysis that may reduce implementation time and help in avoiding common pitfalls. We describe the experiences of a group of academic researchers in using an extensive seven-year repository of US medical.

CHAPTER – 3

METHODS & ALGORITHM USED

3.1 PROJECT DEVELOPMENT:

In this section, there is an overview of the process of data mining and data modelling, from collecting the data, through the data preparation and finally the data modeling. Data cleaning and formatting can be considered as the most critical part of the whole project according to several data scientists. Figure 3.1 shows how the process of data mining works to extract knowledge using algorithms.

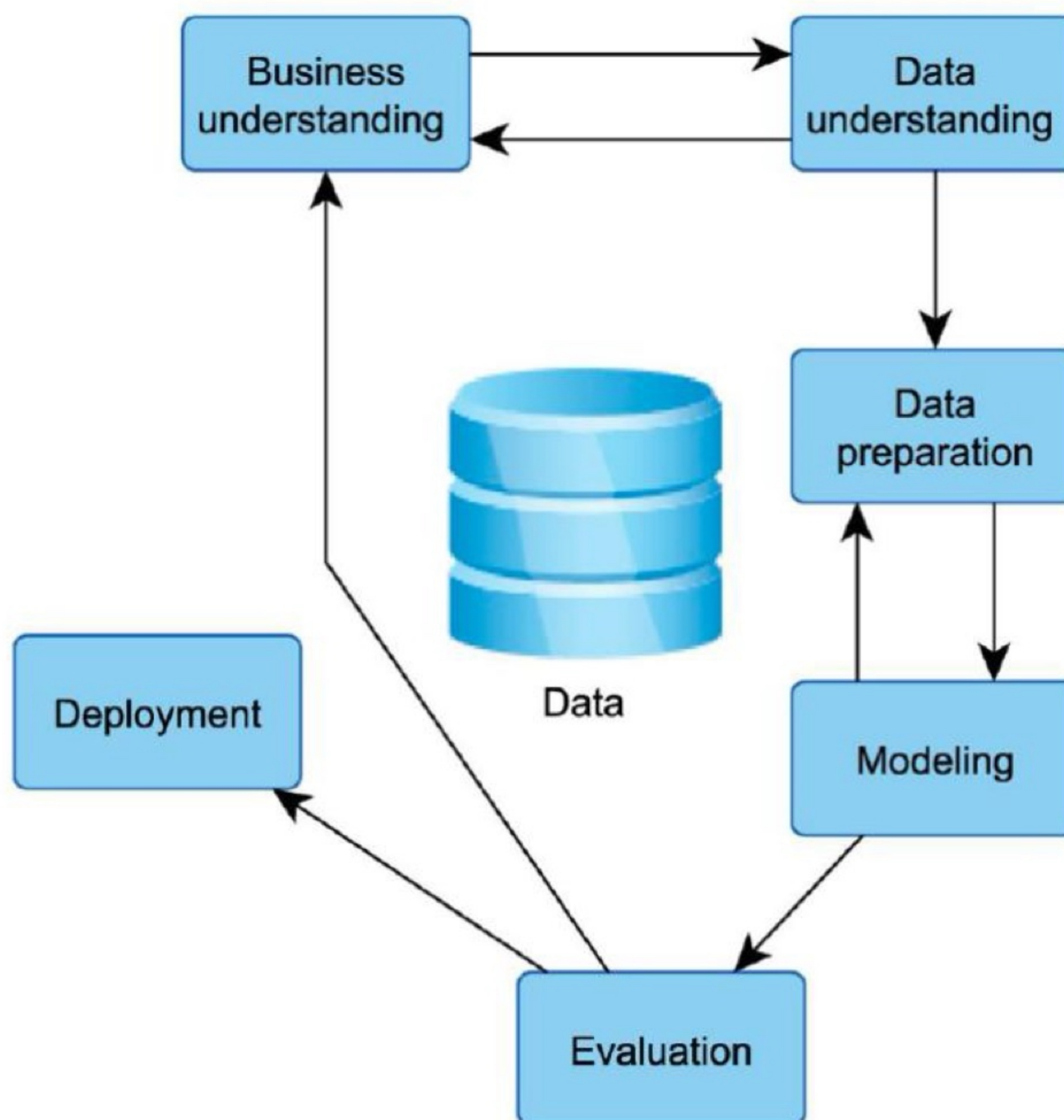


Fig:3.1 Project Development

3.2 DATA ANALYTICS:

The process of data analysis, or alternately, data analysis steps, involves gathering all the information, processing it, exploring the data, and using it to

find patterns and other insights. The process of data analysis consists of:
Data Requirement Gathering: Ask yourself why you're doing this analysis, what type of data you want to use, and what data you plan to analyse.

3.2.1 Data Collection:

Guided by your identified requirements, it's time to collect the data from your sources. Sources include case studies, surveys, interviews, questionnaires, direct observation, and focus groups. Make sure to organize the collected data for analysis.

3.2.2 Data Cleaning:

Not all the data you collect will be useful, so it's time to clean it up. This process is where you remove white spaces, duplicate records, and basic errors. Data cleaning is mandatory before sending the information on for analysis.

3.2.3. Data Analysis:

Here is where you use data analysis software and other tools to help you interpret and understand the data and arrive at conclusions. Data analysis tools include Excel, Python, R, Looker, Rapid Miner, Charito, Meta base, Rehash, and Microsoft Power BI.

3.2.4. Data Interpretation:

Now that you have your results, you need to interpret them and come up with the best courses of action based on your findings.

3.2.5. Data Visualization:

Data visualization is a fancy way of saying, "graphically show your information in a way that people can read and understand it." You can use charts, graphs, maps,

bullet points, or a host of other methods. Visualization helps you derive valuable insights by helping you compare datasets and observe relationships.

3.3 Types of Data Analysis:

The four types of data analysis are:

1. Descriptive Analysis
2. Diagnostic Analysis
3. Predictive Analysis
4. Prescriptive Analysis

3.3.1. DESCRIPTIVE ANALYSIS:

Descriptive analytics answer the question, “What happened?”. This type of analytics is by far the most commonly used by customers, providing reporting and analysis centered on past events. It helps companies understand things such as:

- How much did we sell as a company?
- What was our overall productivity?
- How many customers churned in the last quarter?

Descriptive analytics is used to understand the overall performance at an aggregate level and is by far the easiest place for a company to start as data tends to be readily available to build reports and applications. It's extremely important to build core competencies first in descriptive analytics before attempting to advance upward in the data analytics maturity model. 23 Core competencies include things such as data modeling fundamentals, communicating data with right visualization and basic dashboard skills. It's likely you've adopted some form of descriptive analytics internally, whether that be static P&L statements, PDF reports, or reporting within an analytics tool. For a true descriptive analytics program to be implemented, the concepts of repeatability and automation of tasks must be top of mind. Repeatability in that a data process is standardized and can be regularly applied with minimal

effort (think a weekly sales report), and automation in that complex task (VLOOKUPS, merging of excel spreadsheets, etc.) are automated—requiring little to no manual intervention.

3.3.2. DIAGNOSTIC ANALYSIS:

Diagnostic analytics, just like descriptive analytics, uses historical data to answer a question. But instead of focusing on “what”, diagnostic analytics addresses the critical question of why an occurrence or anomaly occurred within your data. Diagnostic analytics also happens to be the most overlooked and skipped step within the analytics maturity model. Anecdotally, I see most customers attempting to go from “what happened” to “what will happen” without ever taking the time to address the “why did it happen” step. This type of analytics helps companies answer questions such as:

- Why did our company sales decrease in the previous quarter?
 - Why are we seeing an increase in customer churn?
 - Why is a specific basket of products vastly outperforming their prior year sales figures?
- Diagnostic analytics tends to be more accessible and fit a wider range of use cases than machine learning/predictive analytics. You might even find that it solves some business problems you earmarked for predictive analytics use cases. Being at the diagnostic analytics phase likely means you’ve adopted a modern analytics tool. Most modern analytics tools contain a variety of search-based, or lightweight artificial intelligence capabilities. 24 These features allow for detailed insights a layer deeper (for example: the Key Drivers visualization in Power BI, or Qlik’s search-based insight functionality). To be clear, these are an effective lightweight means to address diagnostic analytics use cases but are not a means to a full-scale implementation. Software vendors like Sisu have built their core business around addressing diagnostic analytics use cases (what they call “augmented analytics”) and are a great bet. Diagnostic analytics is an important step in the maturity model that unfortunately tends to get skipped or obscured. If you cannot infer why your sales decreased 20% in 2020, then jumping to predictive analytics and trying to answer “what will happen to sales in 2021” is a stretch in advancing upward in the analytics maturity model.

3.3.3. PREDICTIVE ANALYSIS:

Predictive analytics is a form of advanced analytics that determines what is likely to happen based on historical data using machine learning. Historical data that comprises the bulk of descriptive and diagnostic analytics is used as the basis of building predictive analytics models. Predictive analytics helps companies address use cases such as:

- Predicting maintenance issues and part breakdown in machines.
- Determining credit risk and identifying potential fraud.
- Predict and avoid customer churn by identifying signs of customer dissatisfaction. At the outset of any predictive analytics build, three core elements need to be established:
 - Identify a problem to solve,
 - Define what is you want to predict, and

- State what you will achieve by doing so. To start you should collect existing data, organize data in a useful way to allow for data modeling, cleanse your data and review overall quality, and finally determine your modeling objective. 25 While modeling takes up the spotlight in predictive analytics, data prep is a crucial step that needs to happen first. This is why organizations with a rock-solid foundation in descriptive and diagnostic analytics are better equipped to handle predictive analytics. Simply put, the time and effort to prep, transform, and ensure data quality for retrospective reporting has already taken place. The groundwork should be relatively well laid to quickly identify and leverage data for the modeling phase. I always encourage customers with well-defined KPIs and business logic in a specific business reporting area (think sales reporting for example) to use that as the first predictive analytics use case. The goal is to derive value quickly, and there is no

better place to start than an area where you know data is well defined and of high quality.

3.3.4. PRESCRIPTIVE ANALYSIS:

Prescriptive analytics is the fourth and final pillar of modern analytics. Prescriptive analytics pertains to true guided analytics where your analytics is prescribing or guiding you toward a specific action to take. It is effectively the merging of descriptive, diagnostic, and predictive analytics to drive decision making. Existing scenarios or conditions (think your current fleet of freight trains) and the ramifications of a decision or occurrence (parts breakdown on the freight trains) are applied to create a guided decision or action for the user to take (proactively buy more parts for preventative maintenance). Prescriptive analytics requires strong competencies in descriptive, diagnostic, and predictive analytics which is why it tends to be found in highly specialized industries (oil and gas, clinical health care, finance, and insurance to name a few) where use cases are well defined. Prescriptive analytics help to address use cases such as:

- Automatic adjustment of product pricing based on anticipated customer demand and external factors.

- Flagging selects employees for additional training based on incident reports in the field.

Prescriptive analytics' primary aim is to take the educated guess or assessment out of data analytics and streamline the decision-making process. Prescriptive analytics is commonly considered the merging of descriptive, diagnostic, and predictive analytics. Getting started isn't so much a step-by-step list but rather the time and effort up front to build your competencies within the analytics maturity curve. Simply put, there is no starting point in prescriptive analytics without the requisite first three pillars of modern analytics being established first. If you're ready for prescriptive analytics, then quantifying your call to action and the underlying criteria will be the first requirement. For example: if the use case is to call corrective action for an employee (i.e. – additional training based on poor performance) then the factors that necessitate this action must be firmly established and the action itself

must be clearly defined. Knowing how each kind of analytics helps you better understand your data and how to use it.

3.4 System Requirements:

3.4.1 Windows:

- Microsoft® Windows® 7/8/10 64-bit
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum,
- 4 GB Recommended 500 MB for IDE + 1.5 GB for Android SDK and emulator system image
- 1280 x 800 minimum screen resolution

3.4.2 Mac:

- Mac® OS X® 10.10 Yosemite or higher, up to 10.14 macOS Mojave
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum,
- 4 GB Recommended 500 MB for IDE + 1.5 GB for Android SDK and emulator system image

3.4.3 Linux:

- GNOME or KDE desktop - Tested on gLinux based on Debian.
- 64-bit distribution capable of running 32-bit applications
- GNU C Library glibc 2.19 or later 17
- 4 GB RAM minimum, 8 GB RAM recommended
- 2 GB of available disk space minimum.

3.5 POWER BI:

Power BI is an interactive data visualization software product developed by Microsoft with a primary focus on business intelligence. It is part of the

Microsoft Power Platform. Power BI is a collection of software services, apps, and connectors that work together to turn unrelated sources of data into coherent, visually immersive, and interactive insights. Data may be input by reading directly from a database, webpage, or structured files such as spreadsheets, CSV, XML, and JSON. Microsoft Power BI is a business intelligence (BI) platform that provides nontechnical business users with tools for aggregating, analyzing, visualizing and sharing data. 27 Power BI's user interface is fairly intuitive for users familiar with Excel, and its deep integration with other Microsoft products makes it a versatile self-service tool that requires little upfront training. Users can download an application for Windows 10, called Power BI Desktop, and native mobile apps for Windows, Android and iOS devices. There is also Power BI Report Server for companies that must maintain their data and reports on premises. That version of Power BI requires a special version of the desktop app -- aptly called Power BI Desktop for Power BI Report Server.

3.5.1 USES OF POWER BI:

Microsoft Power BI is used to find insights within an organization's data. Power BI can help connect disparate data sets, transform and clean the data into a data model and create charts or graphs to provide visuals of the data. All of this can be shared with other Power BI users within the organization. The data models created from Power BI can be used in several ways for organizations, including the following:

- telling stories through charts and data visualizations;
- examining "what if" scenarios within the data;
- creating reports that can answer questions in real time and help with forecasting to make sure departments meet business metrics. Power BI can also provide executive dashboards for administrators or managers, giving management more insight into how departments are doing

3.6 METHODOLOGY:

The project is implemented in Microsoft excel and Microsoft power Bi.

- We need to collect the data set for our Insurance Claims, over 2 years or 732 days.
- So, we are creating a data set for our insurance claims
- That includes:
 1. Claims Approved/Denied
 2. Procedure Wise
 3. Age
 4. Gender
 5. Branch Wise
 6. Payer Wise
 7. Amount Wise
 8. Time Taken
 9. Annual Income
 10. Readmission

Now enter the dataset in Microsoft excel sheet as per the date and month as show in the fig 3.2

EXCEL DATA

FileHomeInsertPage LayoutFormulasDataReviewViewHelpTell me what you want to do

Paste

Clipboard

Calibri11A⁻A⁺

B*I*U

Font

Alignment

Number

Styles

Cells

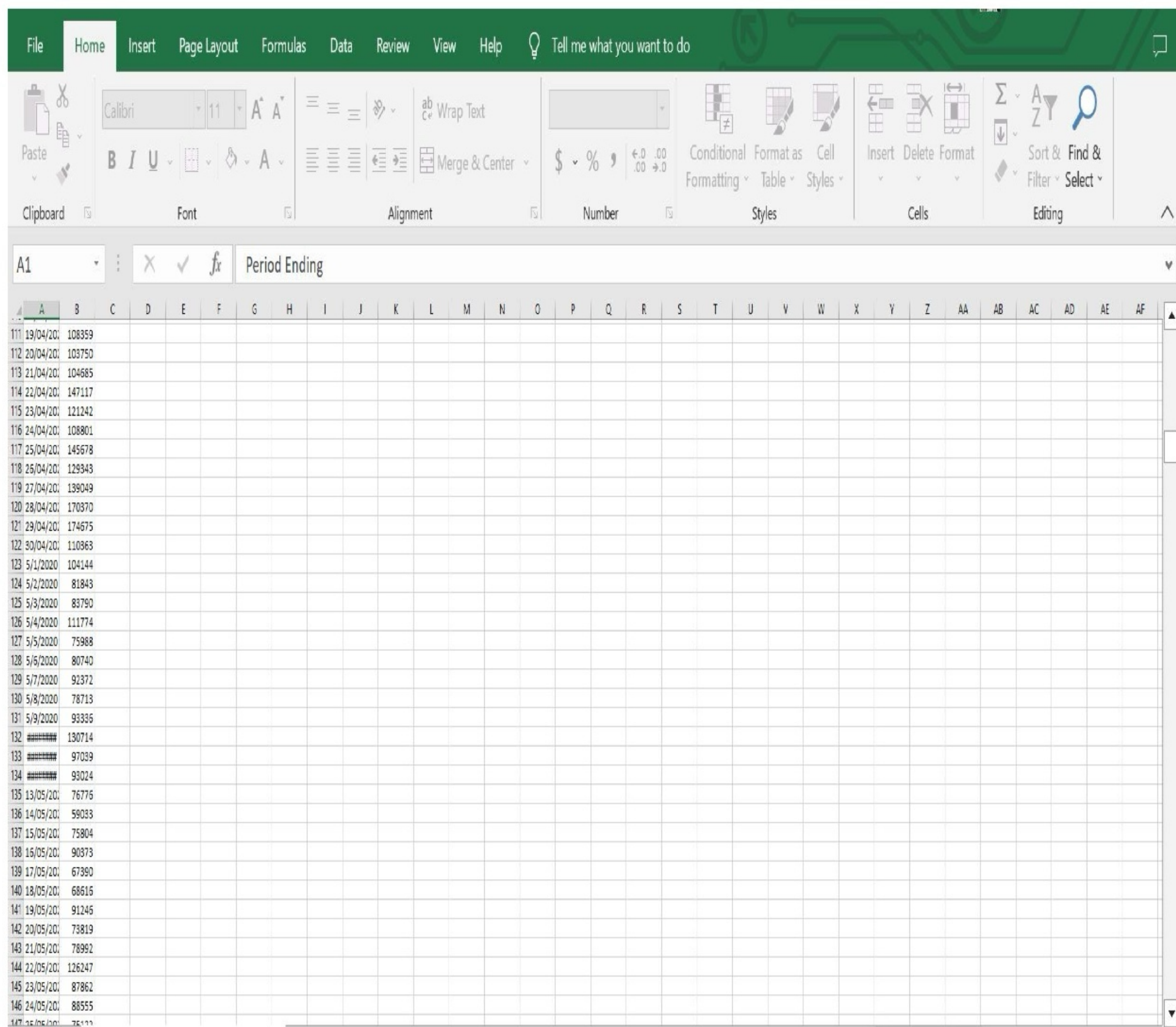
Editing

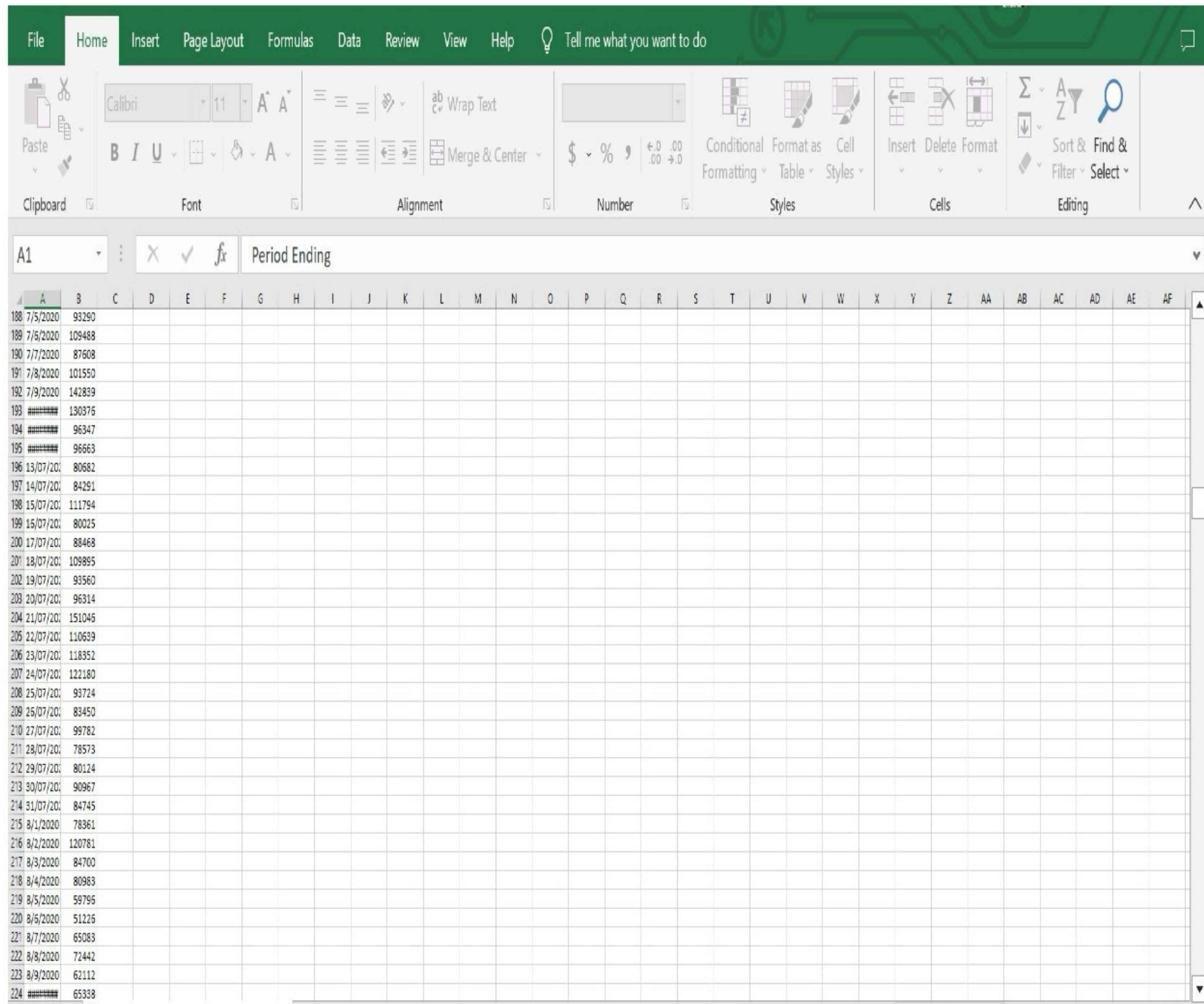
A1

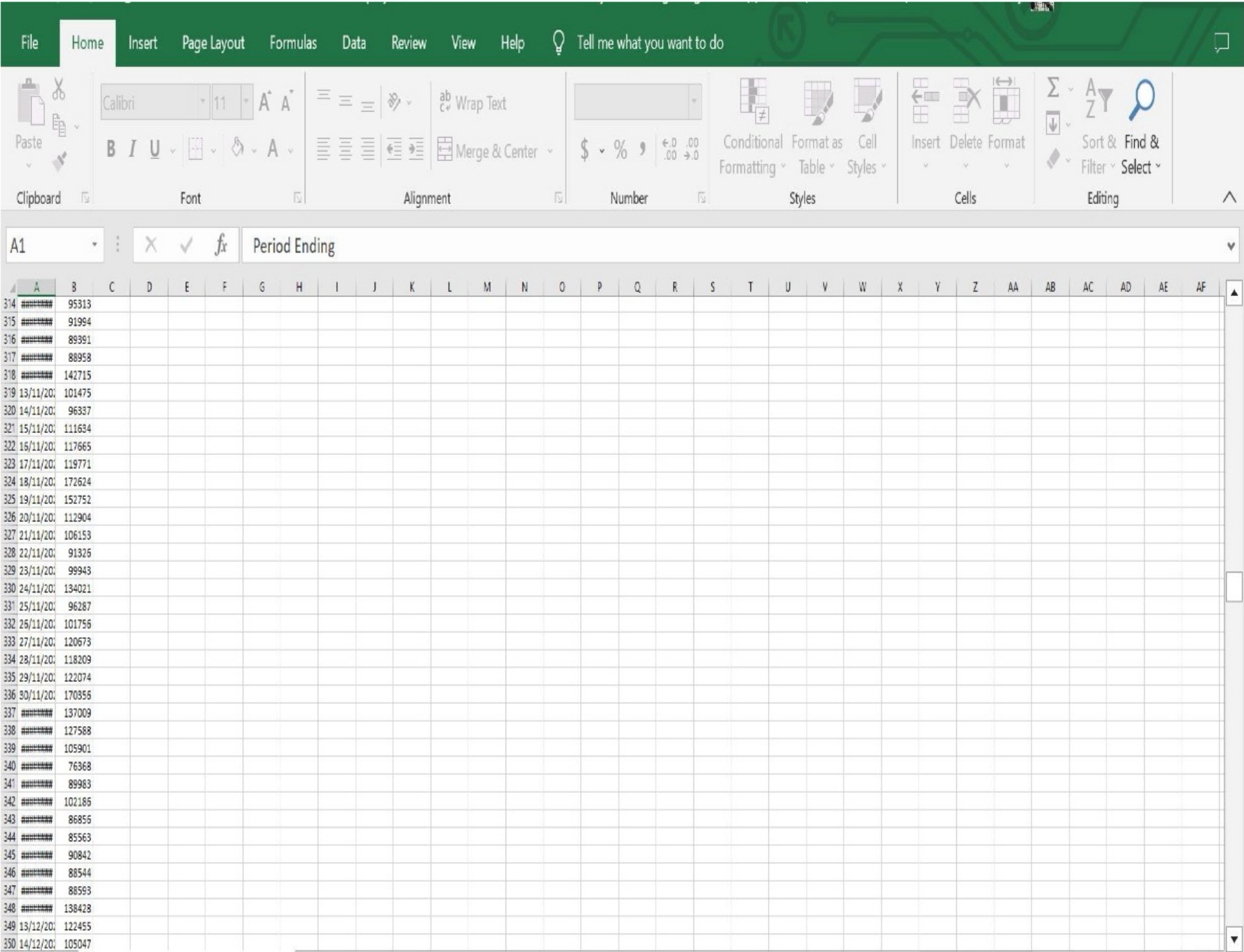
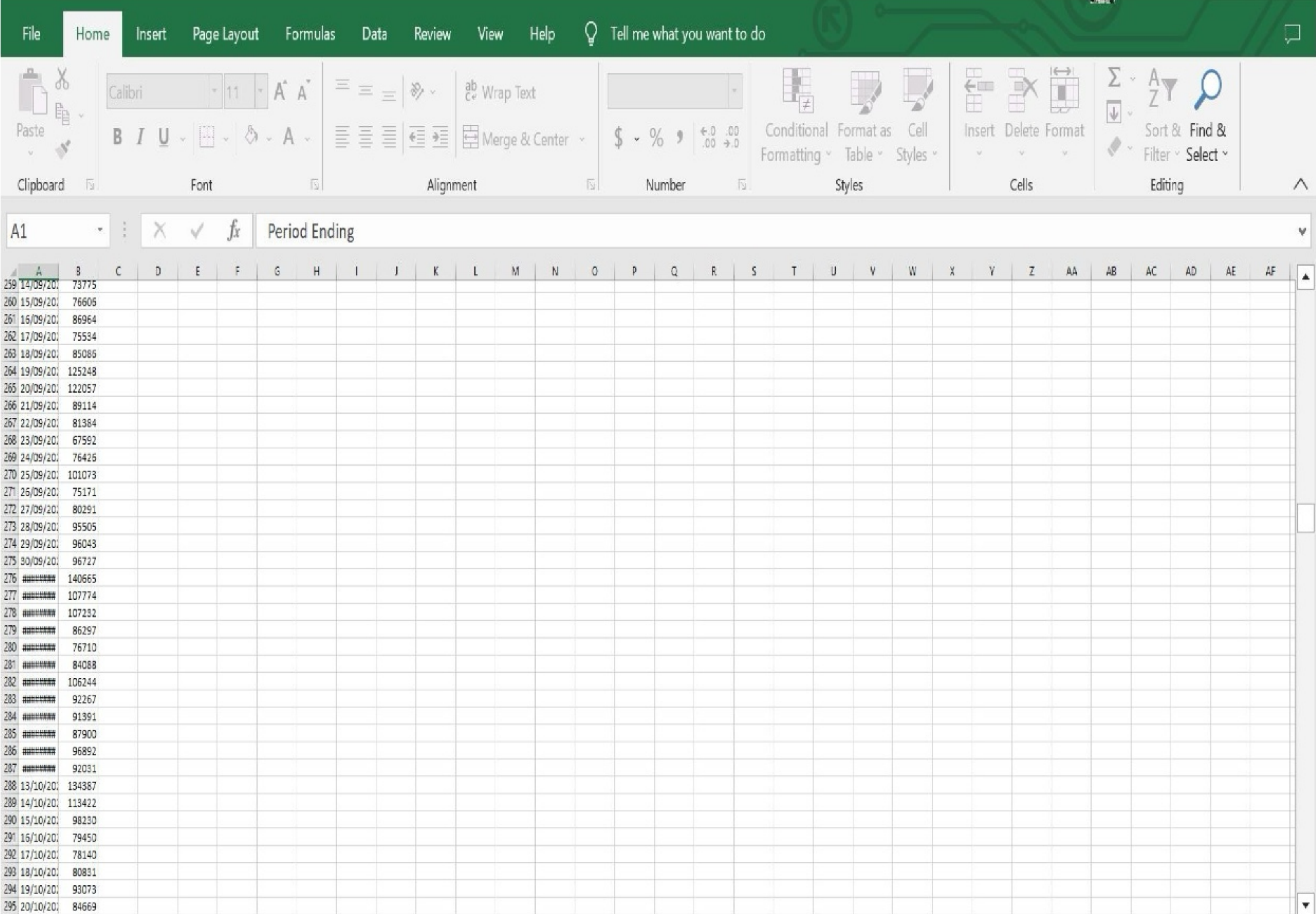
Period Ending

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF
1	Period Ending Initial Claims																															
2	1/1/2020	74556																														
3	1/2/2020	56725																														
4	1/3/2020	49263																														
5	1/4/2020	60250																														
6	1/5/2020	82935																														
7	1/6/2020	55338																														
8	1/7/2020	54965																														
9	1/8/2020	80741																														
10	1/9/2020	58281																														
11	#####	75880																														
12	#####	98931																														
13	#####	98767																														
14	13/01/20:	79057																														
15	14/01/20:	64334																														
16	15/01/20:	56727																														
17	16/01/20:	68873																														
18	17/01/20:	87399																														
19	18/01/20:	65032																														
20	19/01/20:	66971																														
21	20/01/20:	83971																														
22	21/01/20:	70979																														
23	22/01/20:	80451																														
24	23/01/20:	114307																														
25	24/01/20:	84427																														
26	25/01/20:	84295																														
27	26/01/20:	70585																														
28	27/01/20:	56154																														
29	28/01/20:	70087																														
30	29/01/20:	85863																														
31	30/01/20:	68622																														
32	31/01/20:	87931																														
33	2/2/2020	69306																														
34	2/2/2020	71973																														
35	2/3/2020	90880																														
36	2/4/2020	110218																														
37	2/5/2020	102965																														

[illegible]







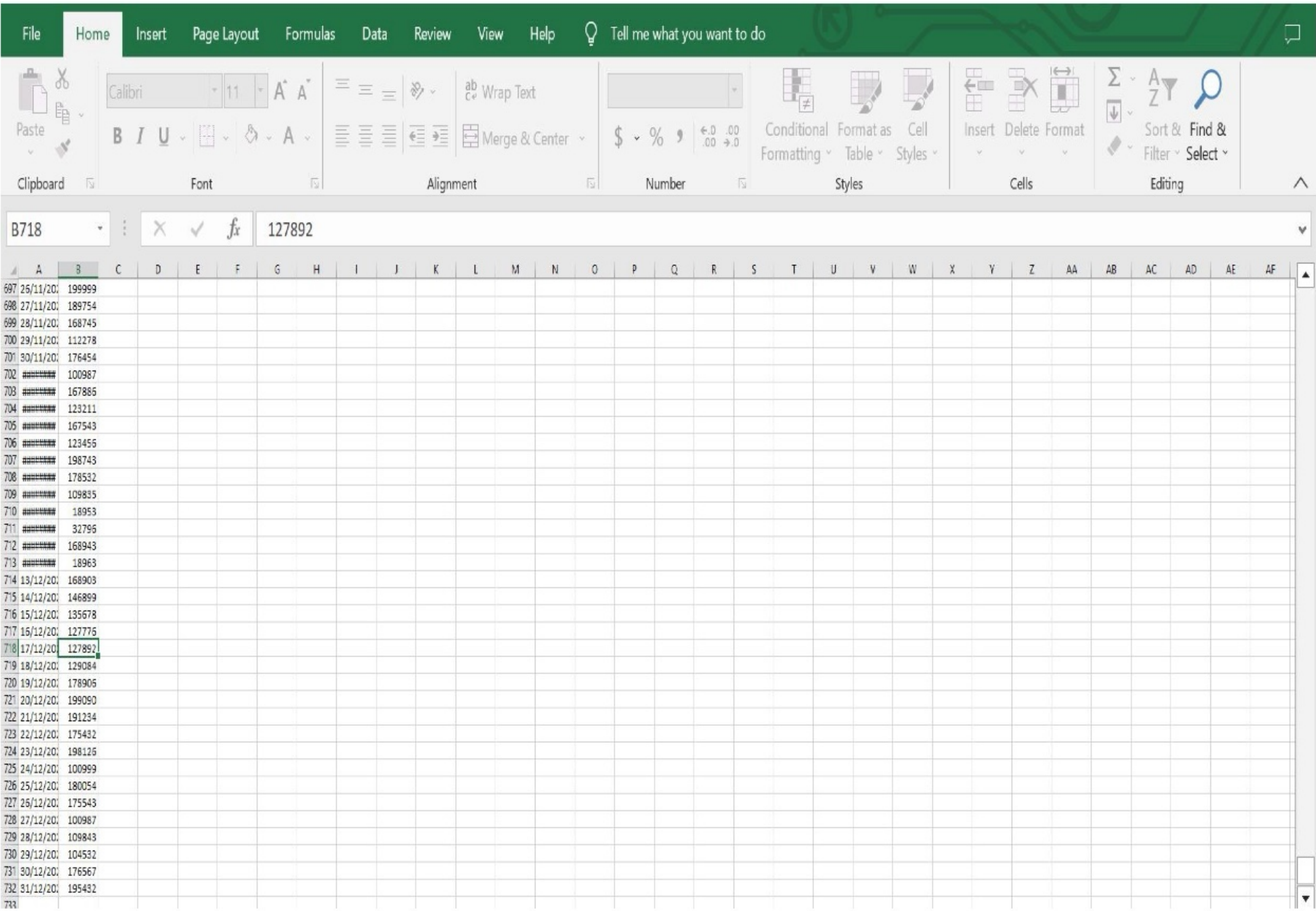
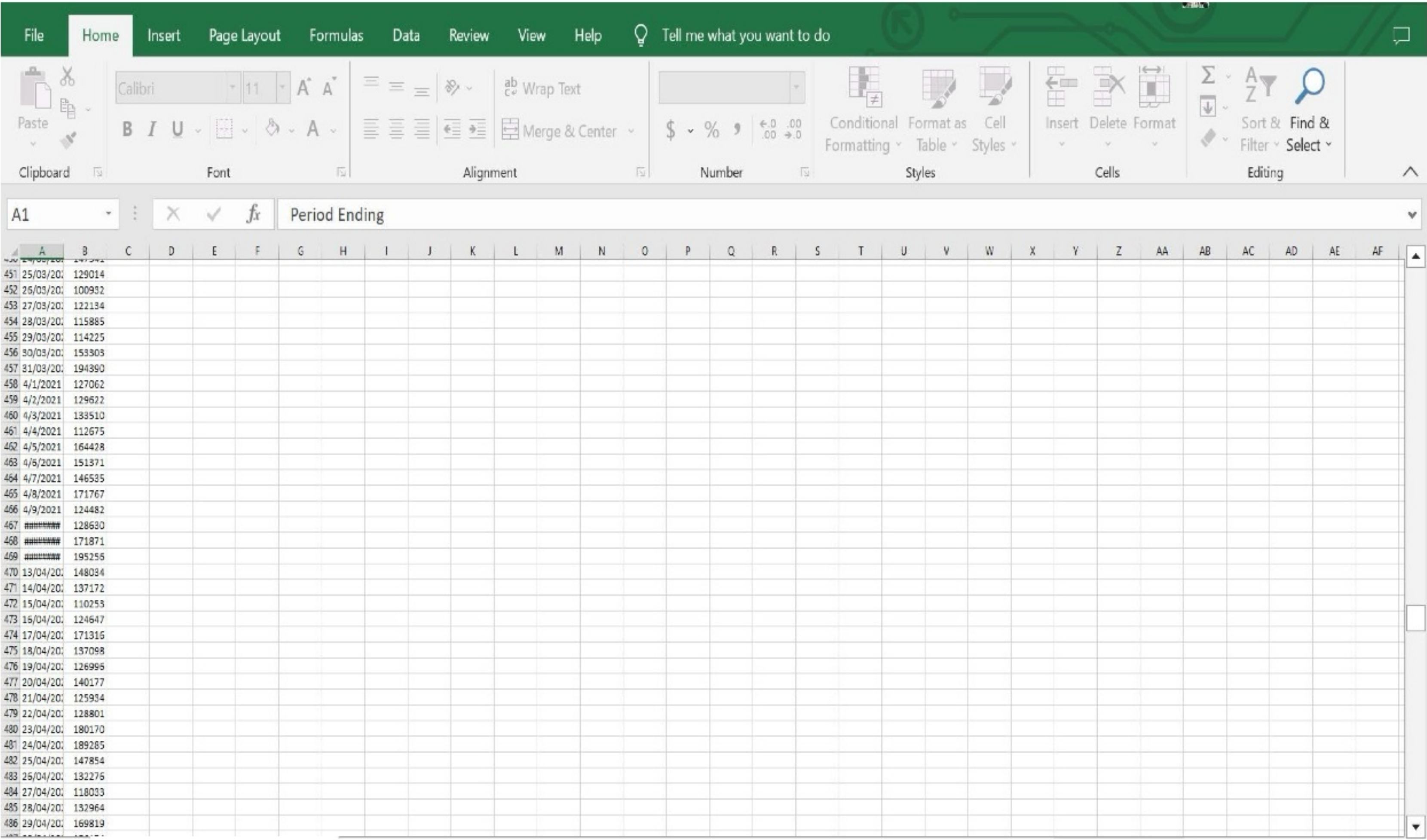


FIG: 3.2 Excel data

- So, we are entering all the data of our claims like, Procedure wise, Gender, Approved/Denied, etc
- Import the excel data in the Power Bi
- Create the required KPI'S
- Here in our project, we created pie chart, bar chart, forecasting etc.
- We also created Q&A Kpi also
- Now we can easily predict the number of claims and claimed amount by using the kpi's
- We can also identify our claimed amount using procedure wise also
- If we click on the bar chart it shows the claimed amount for that procedure.

Claimed Amount by Procedure Wise

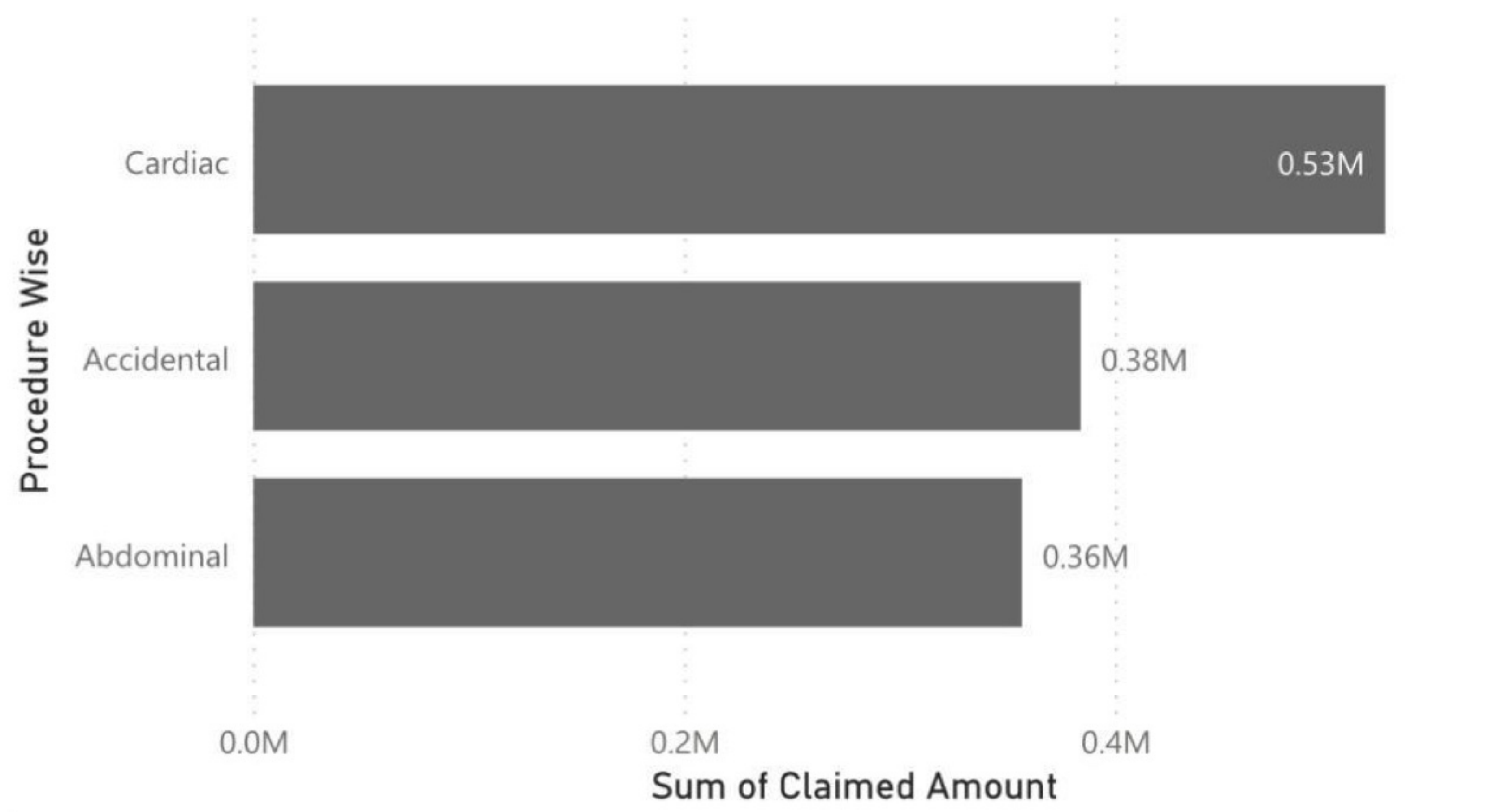


FIG: 3.3 Claimed Amount by Procedure Wise

Power bi makes our work easier so we can recognize the claimed amount on that particular Procedure.

- This bar graph shows Branch wise Claimed Amount.

Here we can see that in the month of Texas, Florida branches are increased as compared to Georgia Branch.

- Same way we can see Georgia branch is the least number of Claimed Amount.

Claimed Amount by Branch Wise

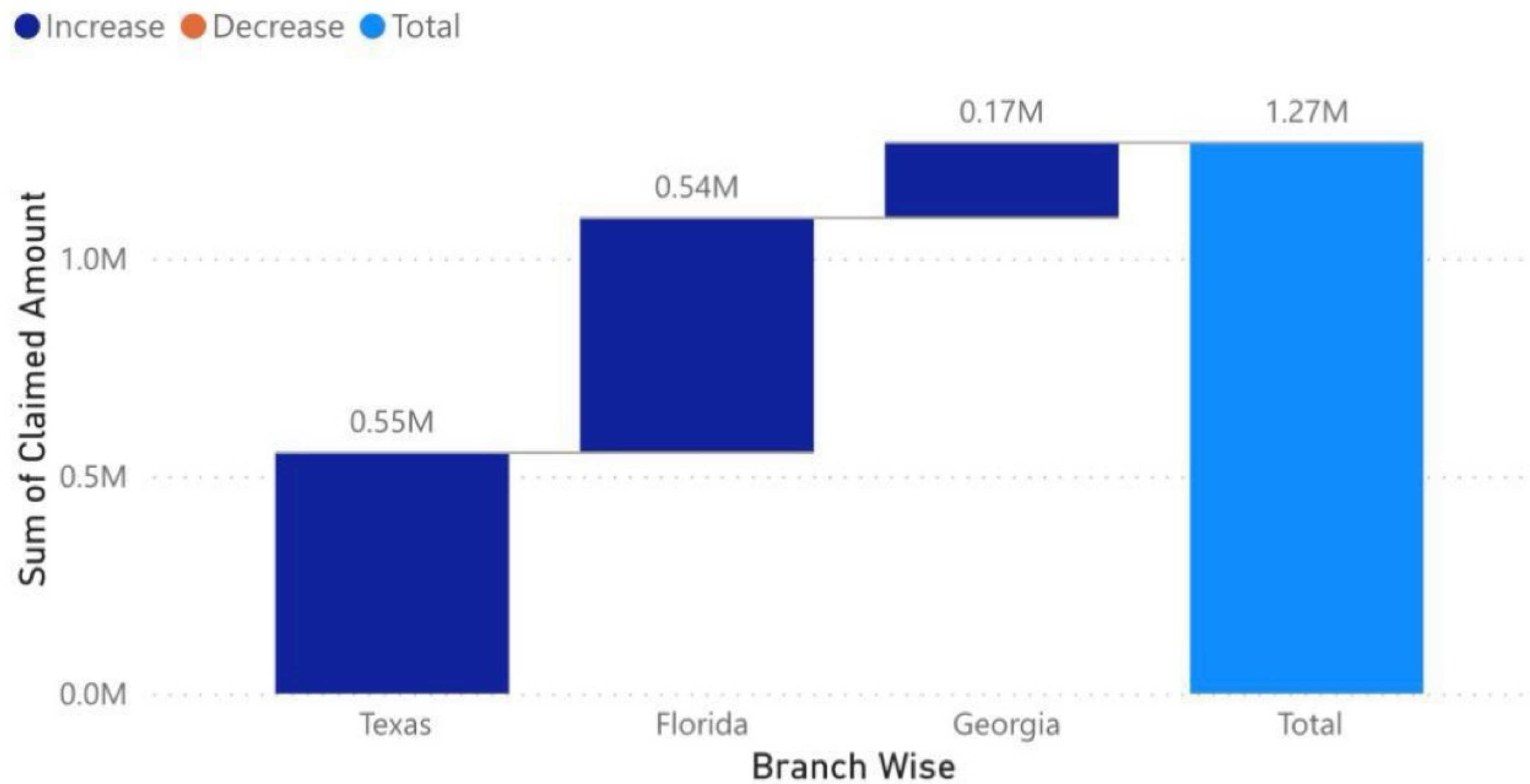


Fig:3.4 Claimed Amount by Branch Wise

- This chart is based on the Claimed Amount by Days taken
- The most claimed amount is 208K in 4 Days
- In other ways the lowest amount is 28K claimed in 6 Days



Fig:3.5 Amount Claimed by Time Taken

- This pie chart represents the Claimed amount by claims Approved/Denied.
- As it seems the claims approved is more than the claims denied.



FIG:3.6 Claims approved/denied

- Here it's a pie chart
- It shows the amount claimed by people of each age
- We can see that older people claim more insurance amount relatively when compared to the younger population.

Claimed Amount by Age

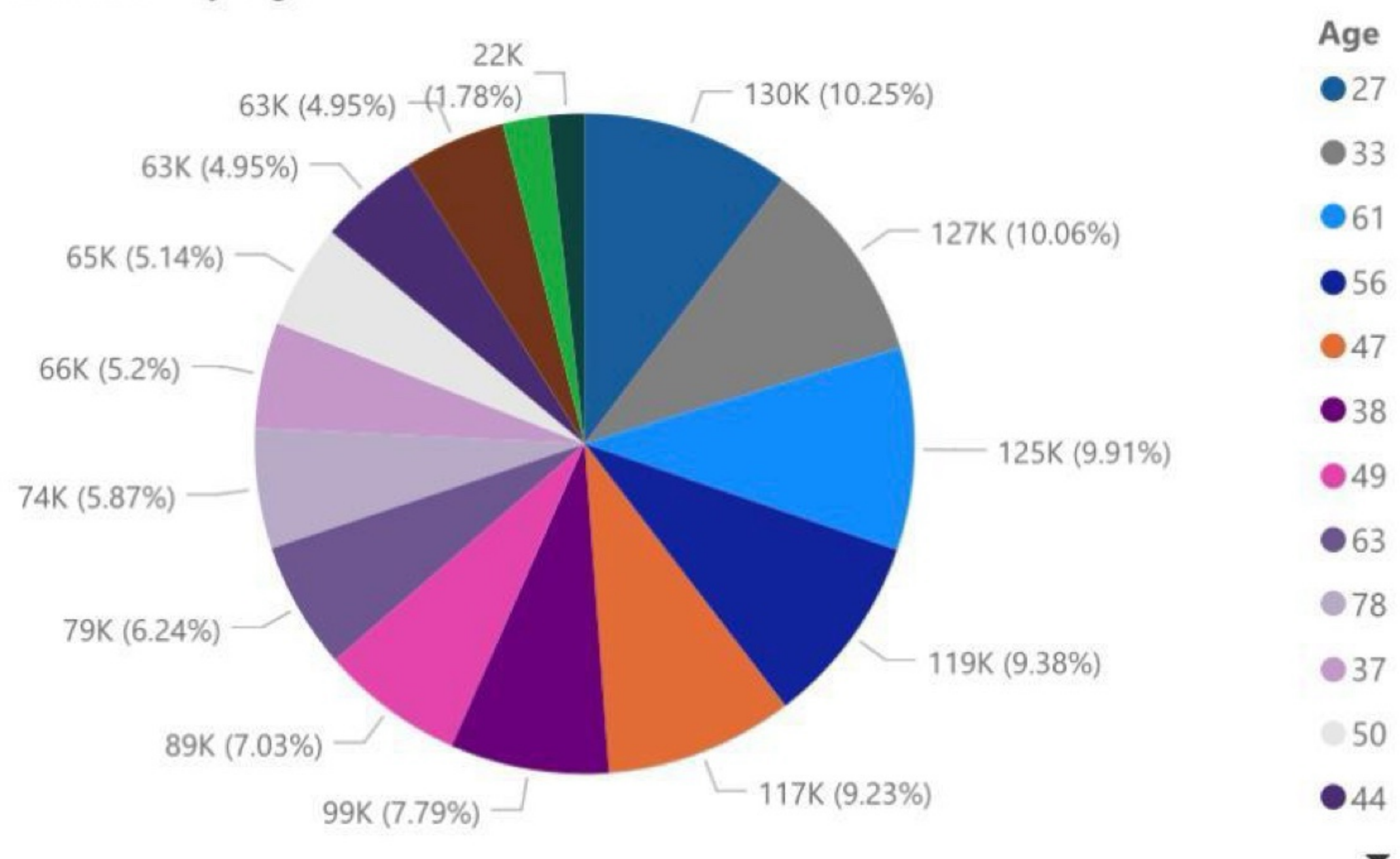


FIG :3.7 Insurance claimed by age

CHAPTER – 4

RESULT & DISCUSSION

The performance of the Prediction is mostly focused on Prediction accuracy, Accuracy in each table and confusion matrix which shows the number of predictions of each table which can be compared to the instances of each table. From the data visualization it was observed that more number of insurance claims and more amount of insurance claims were made during the beginning of 2020 and extended till the end of 2021 which was when COVID outbreak was at its peak. So it can be concluded that the COVID-19 outbreak resulted in a significant increase in the amount of insurance claims made.

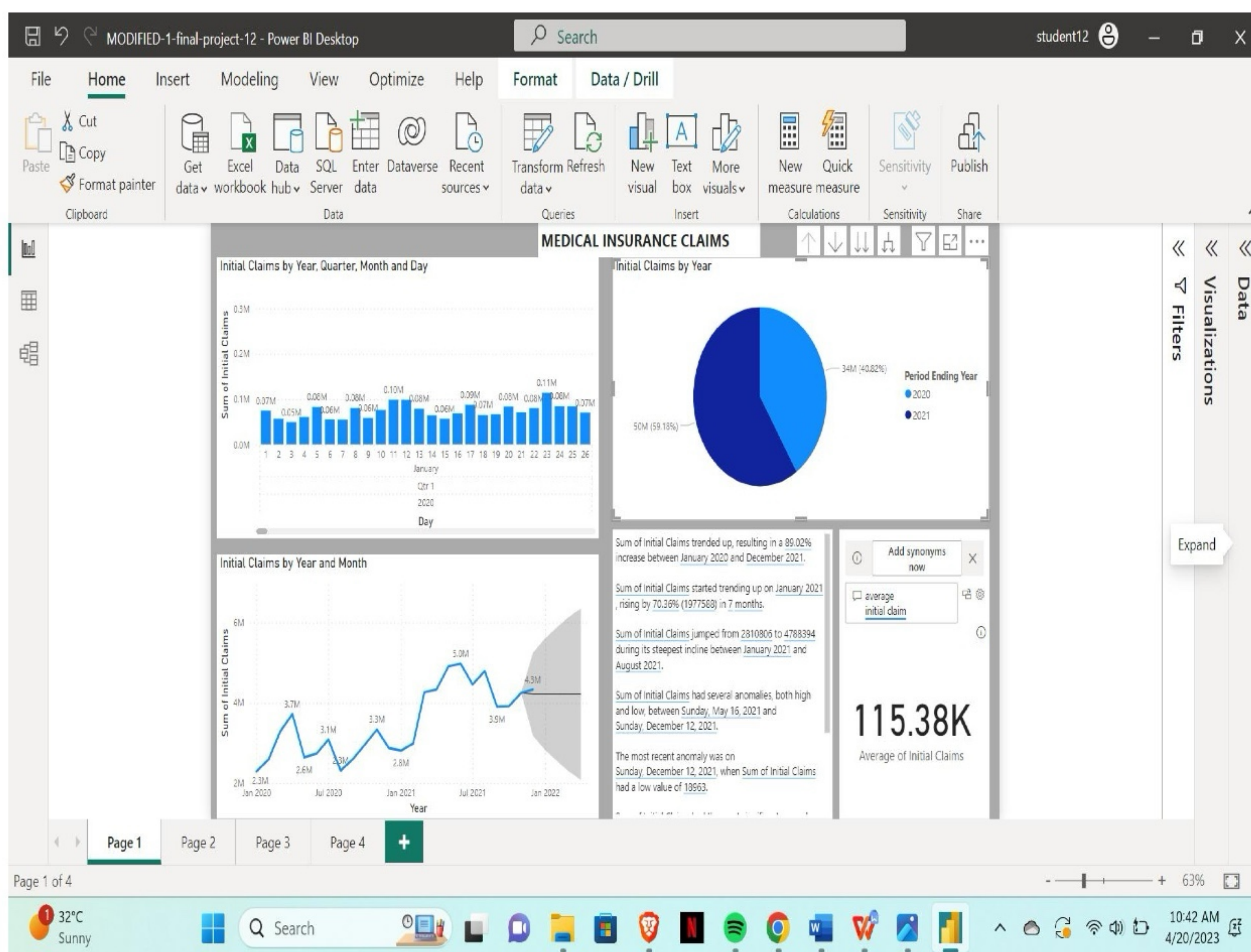
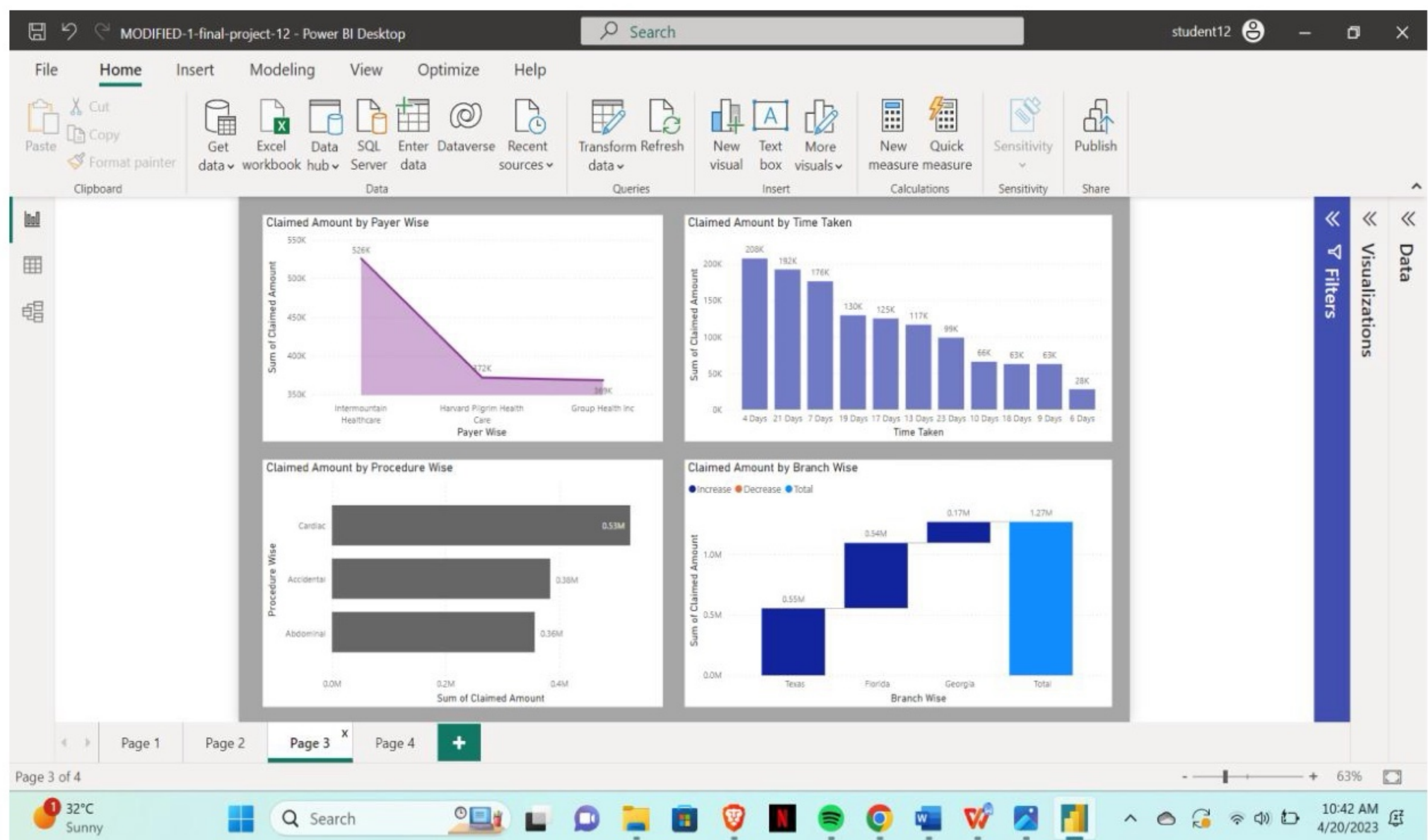
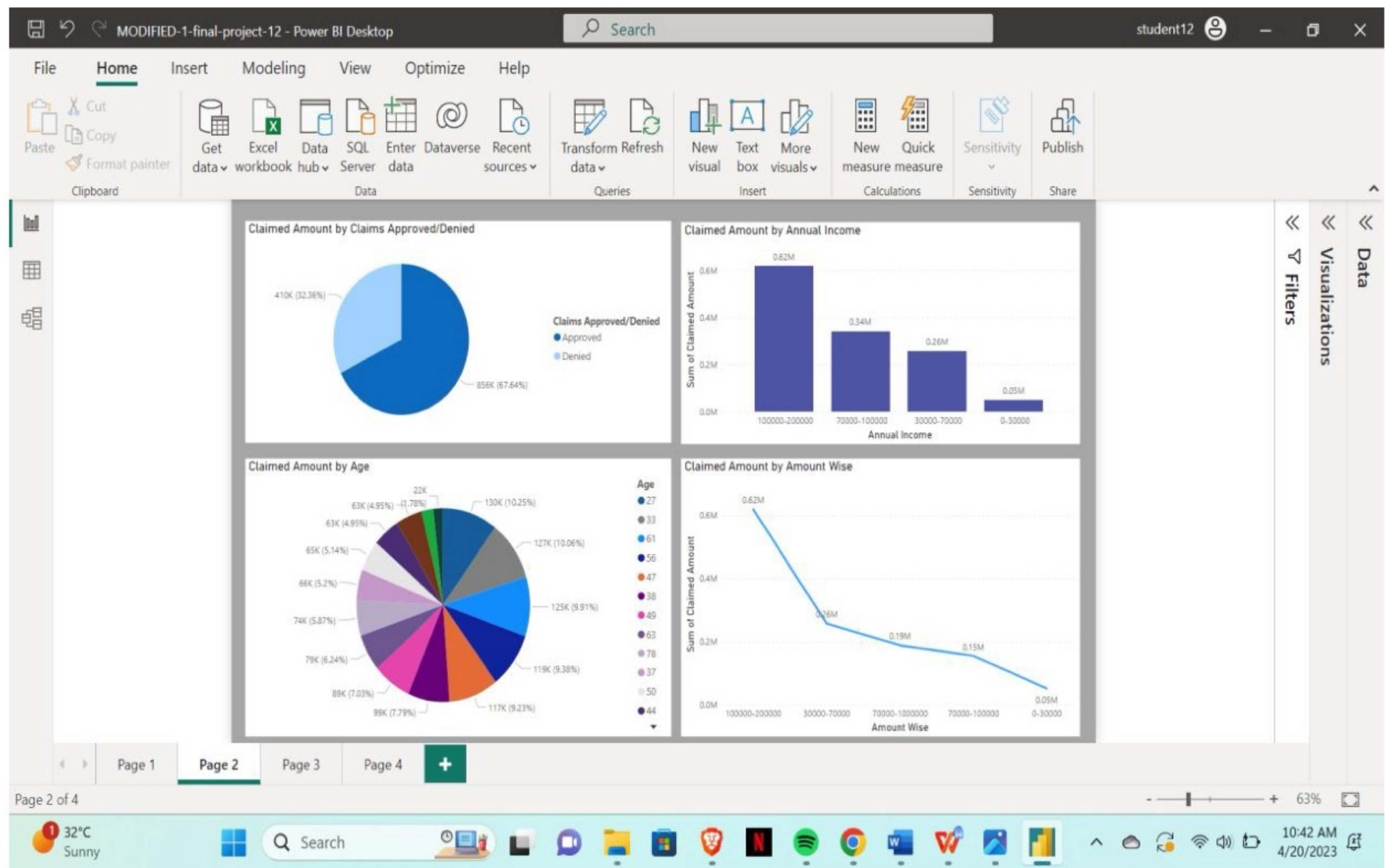
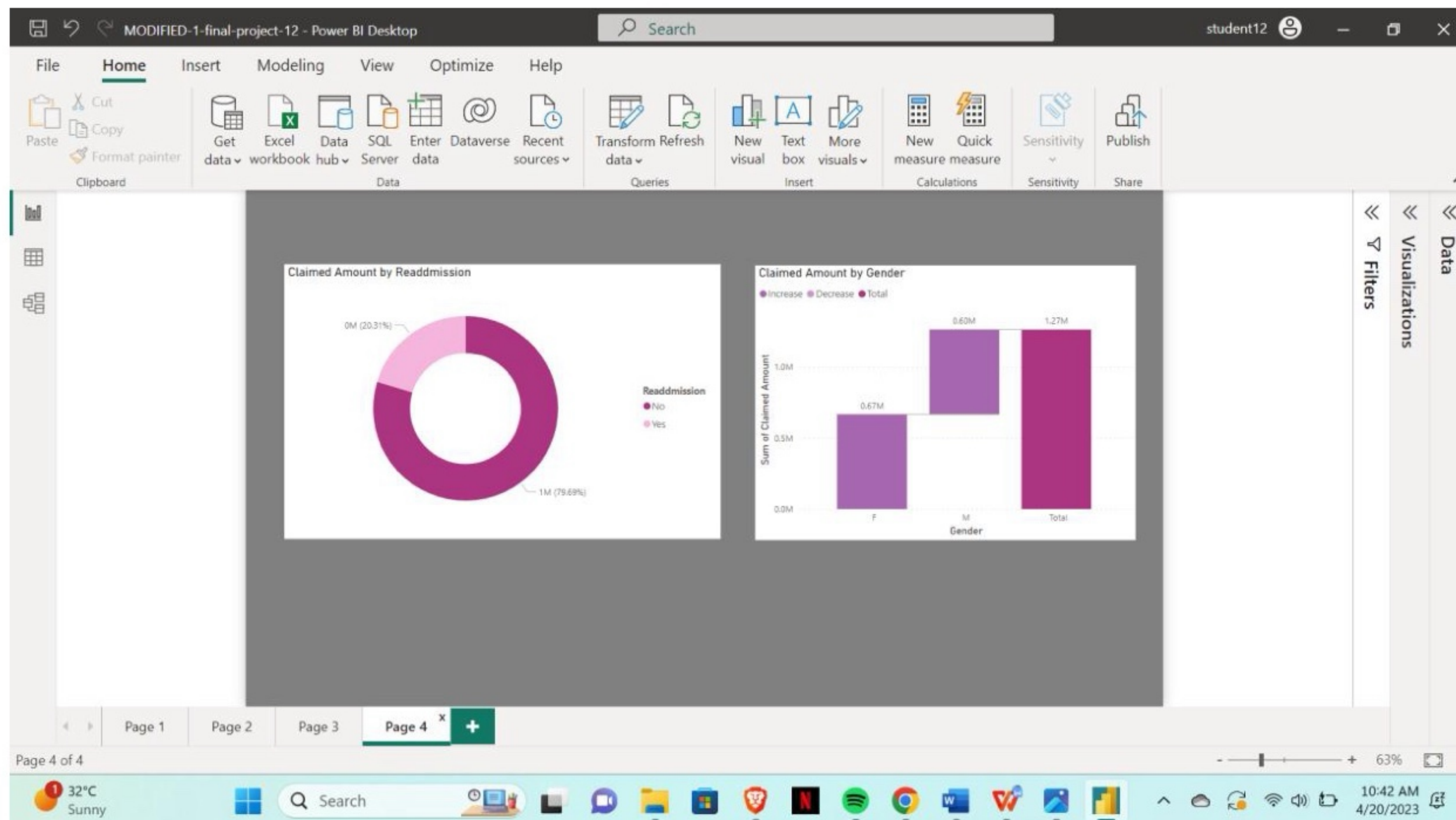


Fig:4.1 MEDICAL INSURANCE CLAIMS IN POWER BI





Medical insurance claims prediction using data analytics is an important field of study in healthcare. The analysis of medical insurance claims data can provide valuable insights into the healthcare industry, including the cost of healthcare services, healthcare utilization patterns, and healthcare outcomes. In this discussion, we will review some of the results and findings from studies that have explored medical insurance claims prediction using data analytics.

Overall, these studies demonstrate the potential of data analytics in predicting medical insurance claims. By using machine learning algorithms to analyze claims data, healthcare providers and insurance companies can identify patients who are at risk of developing chronic diseases, predict hospital readmissions, and accurately predict the cost of medical procedures. These insights can help healthcare providers and insurance companies to develop more effective healthcare policies, improve patient outcomes, and reduce healthcare costs.

CHAPTER – 5

CONCLUSION & FURTHER WORK

5.1 CONCLUSION

The COVID outbreak has caused a significant increase in the number of medical insurance claims and also the amount for which the claims have been made. It can also be concluded that the elderly are more likely to make insurance claims because of the health decline that occurs with age. More insurance claims are made for preplanned surgical procedures than for procedures required after accidents. It is also seen that readmission is very less and most amount is claimed by procedures taking less than 10 days.

5.2 FUTURE ENHANCEMENT

For future scope and improvement of this model All the predicted statistical analysis will be help for proper planning. Data analytics can be used to provide more accurate claims predictions by analyzing historical data and identifying patterns that can be used to forecast future claims. This can help companies, hospitals to make more informed decisions about management, insurance planning, and claims forecasting.

REFERENCES

- Fowles JB, Fowler E, Craft C, et al. Comparing claims data and self-reported data with the medical record for Pap smear rates. *Eval Health Prof.* 1997;20:324-342.
- Motheral B, Brooks J, Clark MA, et al. A checklist for retrospective database studies—report of the ISPOR Task Force on Retrospective Databases. *Value Health.* 2003;6(2):90-97.
- Use of Claims Data Systems to Evaluate Health Care Outcomes Mortality and Reoperation Following Prostatectomy John E. Wennberg, MD; Noralou Roos, PhD; Loreda Sola; et al.
- Use of Insurance Claims Data in Measuring Quality of Care Published online by Cambridge University Press: 10 March 2009 Kathleen N. Lohr
- Estimation of Prevalence of Kidney Disease Treated With Dialysis in China: A Study of Insurance Claims Data, Author open overlay panelChao Yang 1, Zhao Yang 2, Jinwei Wang 1, Huai-Yu Wang 3, Zaiming Su 3, Rui Chen 1, Xiaoyu Sun 3, Bixia Gao 1, Fang Wang 1, Luxia Zhang 1 3, Bin Jiang 4, Ming-Hui Zhao