Communication Data Analysis Findings

Overview of Communication and Sharing Data Analysis Findings

Data Analysis Process

- The data analysis process begins with understanding the problem that needs to be solved and the desired outcome that needs to be achieved. And it ends with communicating the findings in ways that impact decision making.
- Data projects are the result of a collaborative effort spread across business functions involving people with multi-disciplinary skills, with the findings being incorporated into a larger business initiative.
- The success of your communication depends on how well others can understand and trust your insights to take further action.
- So, as data analysts, you need to tell the story with your data by visualizing the insights clearly and creating a structured narrative explicitly targeted at your audience.

Who is your audience?

- Before you begin to create the communication, you need to reconnect with your audience. Begin by asking yourself these questions—Who is my audience? What is important to them? What will help them trust me? Your audience is mostly going to be a diverse group—in terms of the business functions they represent, whether they play an operational or strategic role in the organization, how impacted are they by the problem, and other such factors.

What is important to them?

- Your presentation needs to be framed around the level of information your audience already has.
- Based on your understanding of the audience, you will decide what, and how much, information is essential to enable a better understanding of your findings. It’s tempting to bring out all the data that you’ve been working with, but you have to consider what pieces are more important to your audience than others.

- A presentation is not a data dump. Facts and figures alone do not influence decisions and move people to action. You have to tell a compelling story. Include only that information as is needed to address the business problem. Too much information will have your audience struggling to understand the point you're making.

What will help them trust me?

- Begin your presentation by demonstrating your understanding of the business problem to your audience.
- It's easy to fall back on the assumption that we all know what we're here for, but reflecting your understanding of the problem that needs to be solved, and the outcome that needs to be achieved, is a great first step in winning their attention and starting with trust.
- Speaking in the language of the organization's business domain is another important factor in building a connect between you and your audience.
- The next step in designing your communication is to structure and organize your presentation for maximum impact.
    - Reference the data you have collected. Remember that the data, the very basis of everything that you are communicating, is like a black box for the audience. If you're unable to establish the credibility of your data, people don't know that they can trust your findings. Share your data sources, hypotheses, and validations.
    - Work towards establishing credibility of your findings along the way – don't gloss over any key assumptions made during the analysis.
    - Organize information into logical categories based on the information you have—do you have both qualitative and quantitative information, for example? Be deliberate in taking a top-down or bottom-up approach in your narrative. Both can be effective—depends on your audience and use case.
    - Be consistent in your approach. It's important to determine what communication formats will be most useful to your audience. Do they need to take away an executive summary, a fact sheet, or a report? How is your audience going to use the information you have presented, that should determine the formats you choose.

    - Insights must be explained in a way that inspires action. If your audience doesn't grasp the significance of your insight or are unconvinced of its utility, the insight will not drive any value.

The Role of Visuals

- A thousand-word essay will not have the same impact as a visual in creating a clear mental image in the minds of your audience.
- A powerful visualization tells a story through the graphical depiction of facts and figures.
- Data visualizations—graphs, charts, diagrams—are a great way to bring data to life. Whether you're showing a comparison, a relationship, distribution, or composition, you have tools that can help you show patterns and conclusions about hypotheses.
- Data has value through the stories that it tells. Your audience must be able to trust you, understand you, and relate to your findings and insights.
    - Establishing credibility of your findings, p
    - resenting the data within a narrative,
    - and supporting it through visual impressions, you can help your audience drive valuable insights.

Viewpoints: Storytelling in Data Analysis

In this video, we will listen to data professionals talk about the role storytelling plays and the life of a data analyst.

What role does storytelling play in Data Analysis?

- The role of storytelling in a data analyst's life cannot be overstated. It's supercritical to get really good at storytelling with data.
- I think humans naturally understand the world through stories. If you're trying to convince anyone to do anything with data, the first thing you have to do is tell a clear, a concise, compelling story.
- I also think it can be really useful for the data analyst to develop a story anytime they're working with a dataset to help themselves better understand the underlying dataset and what it's doing.

- There's always going to be a balance between telling a clear, coherent, simple story, and making sure you're conveying all the complexities that you might find within the data. I think finding that balance can be really challenging, but it is really critical.

- The art of storytelling is significant in the life of a data analyst. It doesn't matter how much or what wonderful information you've come up with. If you can't find a way to communicate that to your audience, whether it's the consumer or a director level or executive level person, then it's for naught.
- You have to find a way to communicate that and it's usually best to do it in a visual or through telling a story, so that they understand how that information can be useful.

- I have to say storytelling is essential skill set. It's like the last mile in delivery. A lot of people can handle the technical side through a short period of training.
- However, the ability to extract value from data and to communicate it is in short supply.
- If you think about a long-term career, I think it's very critical to know how to tell a compelling story with data.

- Storytelling is absolutely crucial to data and analytics. This is how you actually convey your message.
- Everyone can show numbers, but if you don't have a story around, if you don't have a compelling reason to act, then ultimately what you're presenting isn't going to resonate with your audience.
- They did a study at Stanford where they had people present their pitches and in that pitch they had simply KPIs, numbers statistics but they also told the story.
- The audience members were then quizzed after the fact, what they remembered from each of those presentations, and it was those stories that stuck with them.
- Yes, there were still facts and figures contained within the story, but that is the way that you drive it home. Having that emotional connection to the story, to the understanding, to

the data is really how you're going to get people to take the action that you want and need them to take.

Introduction to Data Visualization

Overview

- Data visualization is the discipline of communicating information through the use of visual elements such as graphs, charts, and maps. Its goal is to make information easy to comprehend, interpret, and retain.
- Imagine having to look through thousands of rows of data to draw interpretations and compare that to a visual representation of that same data summarizing the findings.
    - Using data visualization, you can provide a summary of the relationships, trends, and patterns hidden in the data, which, if not impossible, would be very hard to decipher from a data dump.

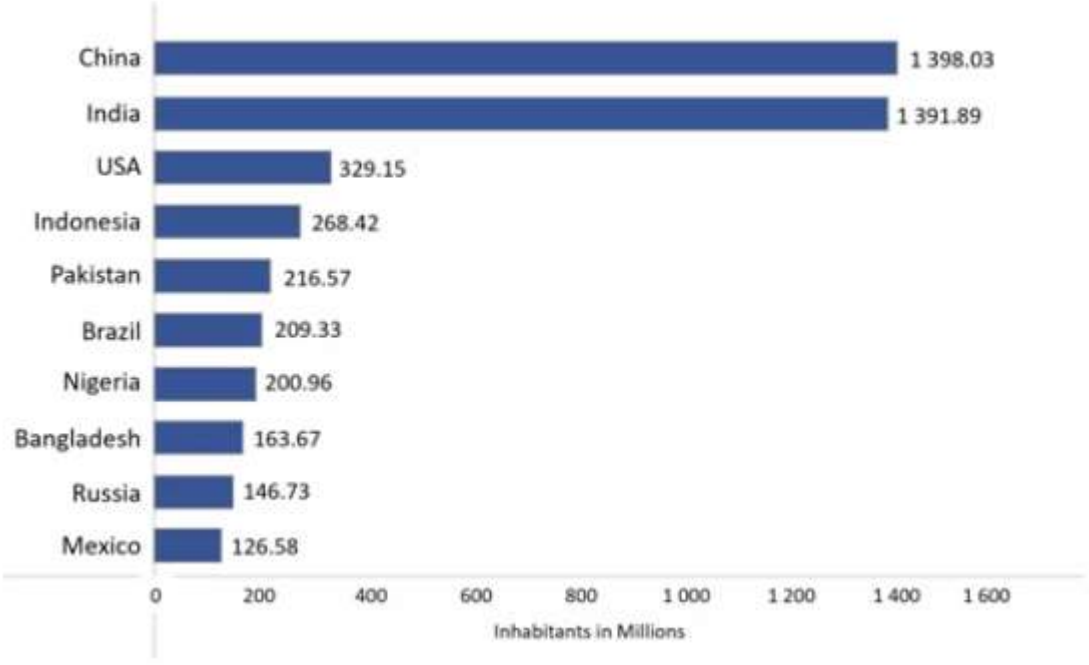
Choosing Appropriate Visualizations

- For data visualization to be of value, you have to choose the visualization that most effectively delivers your findings to your audience.
- And for that, you need to begin by asking yourself some questions.
- What is the relationship that I am trying to establish?
    - Do I want to compare the relative proportion of the sub-parts of a whole, for example, the contribution of different product lines in the total revenue of the company?
- Do I want to compare multiple values, such as the number of products sold, and revenues generated over the last three years? Or, do I want to analyze a single value over time, which in this example could mean how the sale of one specific product has changed over the last three years.
- Do I need my audience to see the correlation between two variables?
    - The correlation between weather conditions and bookings in a ski resort, for example.
- Do I want to detect anomalies in data—for example, finding values in data that could potentially skew the findings?

- What is the question I'm trying to answer is not just an overarching question in the data visualization design and process—you need to be able to answer this question for your audience with every dataset and information that you visualize.
- You also need to consider whether the visualization needs to be static or interactive.
- An interactive visualization, for example, can allow you to change values and see the effects on a related variable in real-time.
- So, think about the key takeaway for your audience,
- anticipate their information needs
- the questions they may have, and then plan the visualization that delivers your message clearly and impactfully.

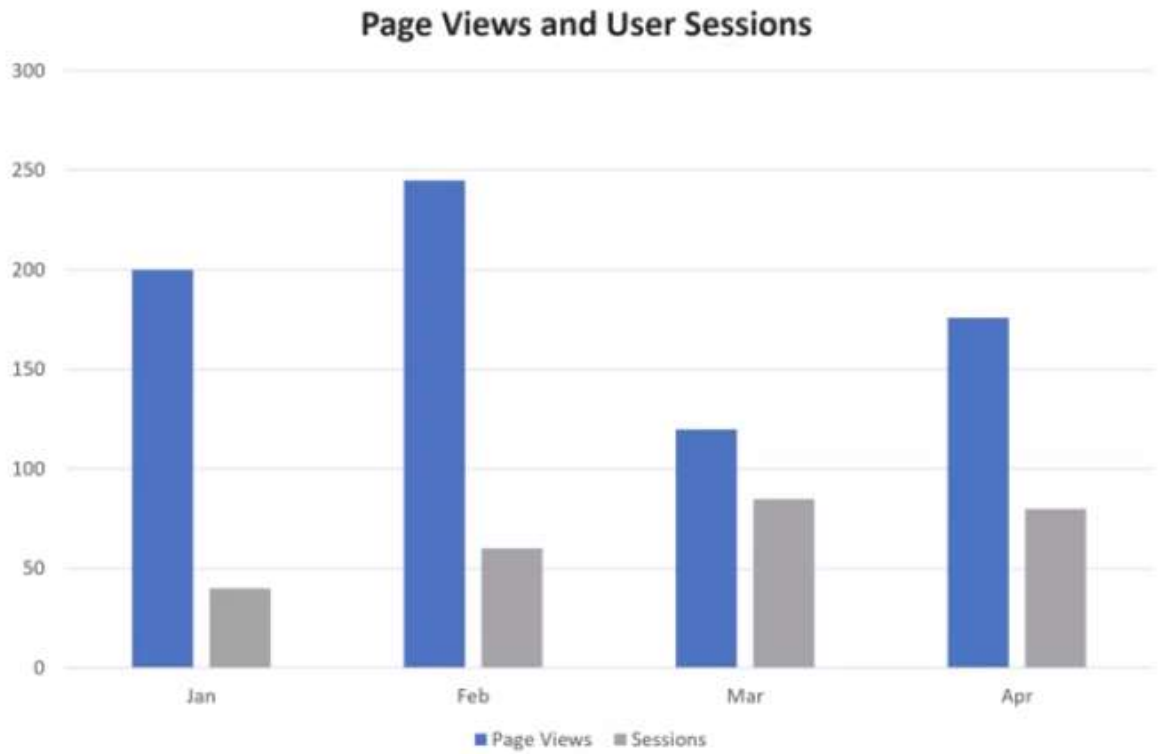
Common types of graphs

- Let's look at some basic examples of the types of graphs you can create for visualizing your data.
- Bar Charts are great for comparing related data sets or parts of a whole.

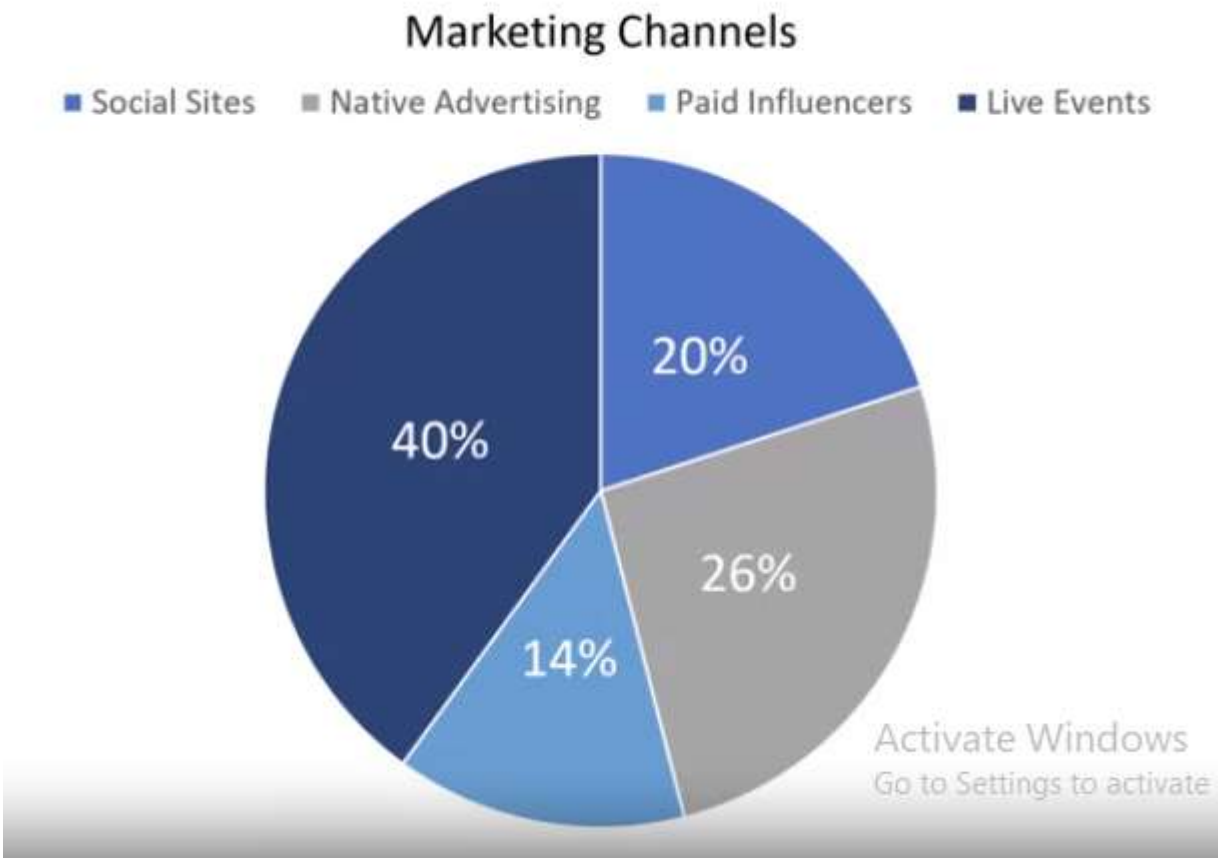


- For example, in this bar chart, you can see the population numbers of 10 different countries and how they compare to one another.

- Column Charts compare values side-by-side. You can use them quite effectively to show change over time.

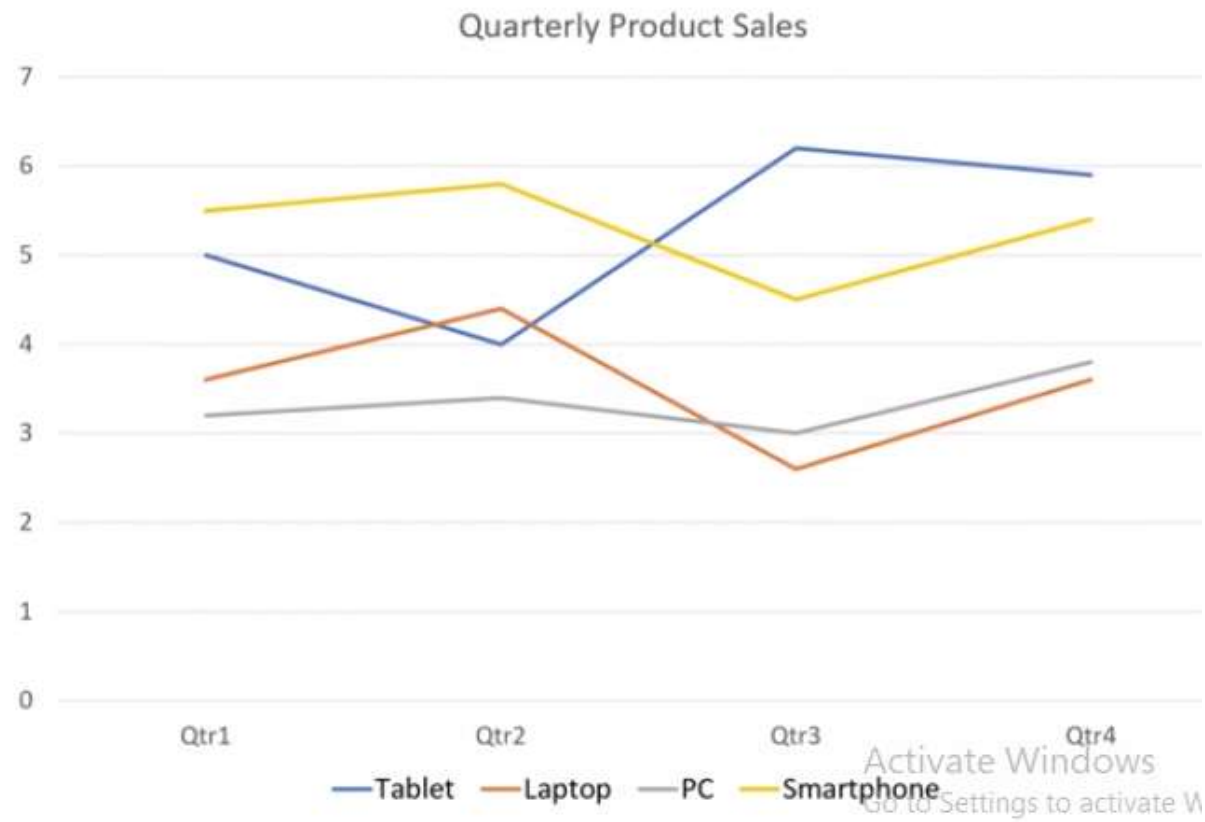


    - For example, showing how page views and user sessions time on your website is changing on a month-to-month basis. Although alike, except for the orientation, bar charts and column charts cannot always be used interchangeably.
    - For example, a column chart may be better suited for showing negative and positive values.
- Pie Charts show the breakdown of an entity into its sub-parts and the proportion of the sub-parts in relation to one another. Each portion of the pie represents a static value or category, and the sum of all categories is equal to hundred percent.



- In this example, in a marketing campaign with four marketing channels—social sites, native advertising, paid influencers, and live events—you can see the total number of leads generated per channel.

- Line Charts display trends. They're great for showing how a data value is changing in relation to a continuous variable.



    - For example, how has the sale of your product, or multiple products, changed over time, where time is the continuous variable.
    - Line charts can be used for understanding trends, patterns, and variations in data; also, for comparing different but related data sets with multiple series.
- Data visualization can also be used to build dashboards. Dashboards organize and display reports and visualizations coming from multiple data sources into a single graphical interface.
    - You can use dashboards to monitor daily progress or the overall health of a business function or even a specific process.
- Dashboards can present both operational and analytical data.
- For example, you could have a marketing dashboard using which you monitor your current marketing campaign for reach-outs, queries generated, and sales conversions, in real-time.
- As part of the same dashboard, you could also be seeing how the conversion rate of this campaign compares to the conversion rate of some of the successfully run campaigns in the past.

- Dashboards are a great tool to present a bird's eye view of the complete picture while also allowing you to drill down into the next level of information for each parameter.
- Dashboards:
    - are easy to comprehend by an average user make collaboration easy between teams; and allow you to generate reports on the go.
    - Using dashboards, you can see the result of variations in data and metrics almost instantly—and this can help you evaluate a situation from multiple perspectives, on the go, without having to go back to the drawing board.

Introduction to Visualization and Dashboarding Software

Overview

- The most commonly used data visualization software and tools.
- These include: Spreadsheets, Jupyter Notebook and Python libraries, R-Studio and R-Shiny, IBM Cognos Analytics, Tableau and Microsoft Power BI.
- Some of these are end-to-end data analytics solutions, while others are specifically for data visualization—ranging from free, open-source tools to commercially available solutions.

Spreadsheets

- Spreadsheets, such as Microsoft Excel and Google Sheets, are possibly the most commonly used software to make graphical representations of data sets.
- Spreadsheets are easy to learn and have a ton of documentation and video tutorials available online for ready reference.
- Excel provides several chart types ranging from the basic bar, line, pie, and pivot charts, to the more advanced options such as scatter charts, trendlines, Gantt charts, waterfall charts, and combination charts (using which you can combine more than one type of charts).
    - Excel also provides recommendations on the best visual representation for your data set.
    - To make the charts more presentable, you can add a chart title, change colors of the elements, and add labels to data.
- Google Sheets also offers similar chart types for visualization, though Excel does have more inbuilt formula-based options than Google Sheets.

    - Like Excel, Google Sheets can help you choose the right visualization. All you have to do is highlight the data you wish to visualize and click the chart button—and you get a list of suggested charts best suited for your data.
    - Charts and reports automatically update, in Excel as well as in Google Sheets, as the underlying data is changed.
    - Google Sheets is preferred over Excel, where multiple users need to collaborate.

Jupyter Notebook and Python Libraries

- Jupyter Notebook is an open-source web application that provides a great way to explore data and create visualizations. You don't have to be a Python expert to use Jupyter Notebook.
- Python provides a host of libraries that are used for data visualization. Let's look at a few of those libraries.
- Matplotlib is a widely used Python data visualization library.
    - It provides different kinds of 2D and 3D plots and the flexibility to create plots in several different ways.
    - Using Matplotlib, you can create high-quality interactive graphs and plots with just a few lines of code.
    - It has large community support and cross-platform support as it is an open-source tool.
- Bokeh provides interactive charts and plots and is known for delivering high-performance interactivity over large or streaming datasets.
    - Bokeh offers flexibility for applying interaction, layouts, and different styling options to visualization.
    - It can also transform visualizations written in some of the other Python libraries, such as Matplotlib, Seaborn, and Ggplot.
- Dash is a Python framework for creating interactive web-based visualizations.
    - Using Dash, you can build highly interactive web applications using Python code.
    - While knowledge of HTML and javascript is useful, but it is not a requirement.
    - Dash is easily maintainable, cross-platform, and mobile-ready.

R-Studio and R-Shine

- Using R-Studio, you can create basic visualizations such as histograms, bar charts, line charts, box plots, and scatter plots; and advanced visualizations such as heat maps, mosaic maps, 3D graphs, and correlograms.
- Shiny is an R package that helps build interactive web apps that you can host as standalone apps on a webpage.
    - These web apps seamlessly display R objects, such as plots and tables, and can be made live to allow access to anyone. You can also build dashboards using Shiny. The ease of working with Shiny is what popularized it among data professionals.

IBM Cognos Analytics

- IBM Cognos Analytics is an end-to-end analytics solution.
- Some of the visualization features provided by Cognos include:
    - Importing custom visualizations;
    - A forecasting feature that provides time-series data modeling and forecasts based on data presented in corresponding visualizations;
    - Recommendation for visualizations based on your data;
    - Conditional formatting which allows you to see the distribution of your data and highlight exceptional data points, for example, highlighting high and low sales numbers over a certain threshold;
    - Cognos is known for its superior visualizations and overlaying data on the physical world using its geospatial capabilities.

Tableau

- Tableau is a software company that produces interactive data visualization products.
    - Using tableau products, you can create interactive graphs and charts in the form of dashboards and worksheets, with drag and drop gestures.
    - Tableau also offers the option to publish results in the form of stories.
    - You can import R and Python scripts in Tableau and take advantage of its visualization features that are far more superior to that of other languages.

- Tableau’s visualization capabilities are easy and intuitive to use. Tableau is compatible with excel files, text files, relational databases, and cloud database sources such as Google Analytics and Amazon Redshift.

Power BI

- Power BI is a cloud-based business analytics service from Microsoft that enables you to create reports and dashboards.
- It is a powerful and flexible tool known for its speed and efficiency, and an easy to use drag and drop interface.
- Power BI is compatible with multiple sources, including Excel, SQL Server, and cloud-based data repositories, which makes it an excellent choice for data professionals. Power BI provides the ability to collaborate and share customized dashboards and interactive reports securely, even on mobiles.
- Power BI’s dashboard consists of many visualizations on a single page that help you tell your story.
- These visualizations, called tiles, are pinned to the dashboard. The dashboard is interactive, which means a change in one tile affects the other.

Considerations

- When deciding which tools to use, you need to consider the ease-of-use and purpose of the visualization. In terms of the tools that are available and the visualization capabilities they offer —if you can visualize it, you can create it.

Viewpoints: Visualization Tools

What are the visualization tools you rely on the most, and why?

- >> The visualization tool that I rely on most of my day to day life is Cognos Analytics.
    - There's a few reasons for this. One, it allows me to very quickly important spreadsheet, connect to a database and visualize my data, whether that's me understanding what I want to look at and dragging the fields on, or using our AI assistant to present the data to help me understand and explore what might be interesting in there, if it's a new data set, I've never worked with before.

- Now, on top of that, I can also go ahead and start to do some more complex things, or even just some more robust analysis with our reporting tool to allow me to build out and schedule reports for delivery. If I wanted my sales seem to have their pipeline report or their sales opportunity report every Monday morning, set that up once, use what we call bursky, and then have that sent out automatically every Sunday night, so it's waiting for them in the morning. On top of this I can search combine multiple data sources and how the system helps me create those joins together. And then be able to visualize those all on a symbol simple dashboard that's highly interactive, allowing you to filter and sort dynamically, as well as share that with the rest of an organization, so that not every user has to go through the same experience. We've set up dashboard once, everyone can then have access to it.
- >> In terms of visualization tools, I rely the most on Looker, which is a data visualization tool that sits atop my company's internal database. It's similar to Tableau, which I've also used in the past and find pretty easy to use. And the great thing about these data visualization tools like Looker and Tableau is they let everyone throughout the organization, regardless of whether or not they're data professionals, easily kind of see their data and do basic aggregation or sorting on it. A data visualization tool I really rely on for exploratory data analysis is R. I've been a big convert in recent years to the effectiveness of doing basic data analysis and data visualization in R, particularly using the Tidyverse, which is a collection of packages that help you really easily load in your data, aggregate it at different levels, and also quickly and easily visualize it.
- >> Tableau and the Power BI no brainer, they're easy to pick up and very helpful to demonstrate data. And as more and more companies and people started to utilize them, there are more and more built in templates in libraries.
- >> I would say the visual, Would probably be Excel and Word, just the Microsoft suite, and just looking at and using the sums and the macros to make sure that the data when I, before I even dive in those coming and then it makes sense and that it is prepped for what we need it to be.

Summary

- In this lesson, you have learned the following information:

- Data has value through the stories that it tells. In order to communicate your findings impactfully, you need to:
    - Ensure that your audience is able to trust you, understand you, and relate to your findings and insights.
    - Establish the credibility of your findings.
    - Present the data within a structured narrative.
    - Support your communication with strong visualizations so that the message is clear and concise, and drives your audience to take action.
- Data visualization is the discipline of communicating information through the use of visual elements such as graphs, charts, and maps. The goal of visualizing data is to make information easy to comprehend, interpret, and retain.
- For data visualization to be of value, you need to:
    - Think about the key takeaway for your audience.
    - Anticipate their information needs and questions, and then plan the visualization that delivers your message clearly and impactfully.
- There are several types of graphs and charts available for you to be able to plot any kind of data, such as bar charts, column charts, pie charts, and line charts.
- You can also use data visualization to build dashboards. Dashboards organize and display reports and visualizations coming from multiple data sources into a single graphical interface. They are easy to comprehend and allow you to generate reports on the go.
- When deciding which tools to use for data visualization, you need to consider the ease-of-use and purpose of the visualization. Some of the popularly used tools include Spreadsheets, Jupyter Notebook, Python libraries, R-Studio and R-Shiny, IBM Cognos Analytics, Tableau, and Power BI.