Visualization and Storytelling

Dhiraj Lal Pradhan
Senior Data Analyst, Fusemachines

What is Data Visualization

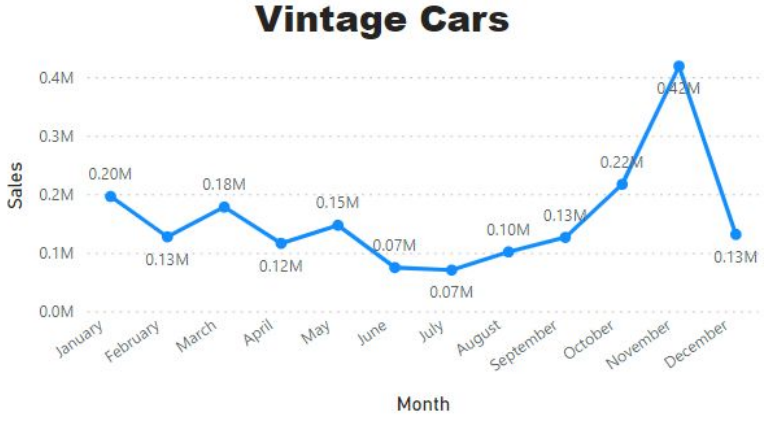
Data visualization is the graphical representation of information and data. It provides an accessible way to see and understand trends, outliers, and patterns in data by using visual elements like charts, graphs, and maps.

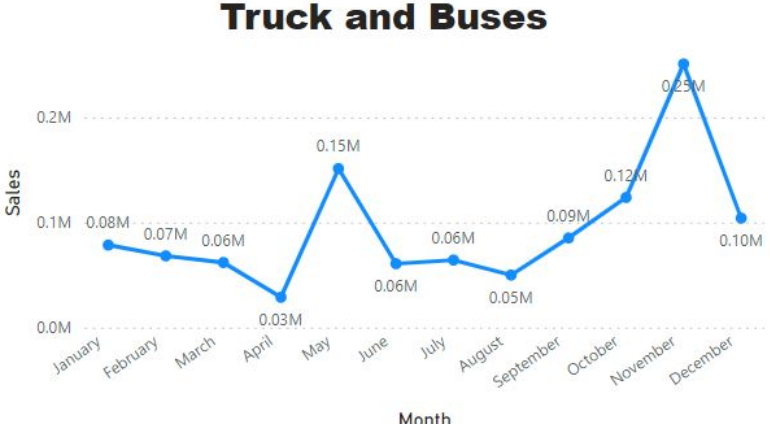
BRINGING YOUR DATA TO LIFE

Sales By Month

Month Name	VIntage Cars	Trucks and Buses	Trains	Motorcycles	Classic Cars
January	196130	78531	19027	81114	303071
February	127120	68211	16508	122802	299648
March	177935	61877	22581	60470	277561
April	115942	28791	4756	119607	263252
May	146877	151270	20457	108336	365947
June	74568	60778	10071	49879	137420
July	70365	64270	10802	60698	245293
August	101425	49943	16168	106870	265302
September	126146	85298	12480	45627	234673
October	216764	123814	25417	103650	465002
November	418664	250874	44795	261057	825156
December	131217	104134	23181	46279	237291
Total	**1903151**	**1127790**	**226243**	**1166388**	**3919616**


Trains
Sales
40K
30K
20K
10K
0K
19K
17K
23K
5K
20K
10K
11K
16K
12K
25K
45K
23K
January
February
March
April
May
June
July
August
September
October
November
December
Month


Vintage Cars
Sales
0.4M
0.3M
0.2M
0.1M
0.0M
0.20M
0.13M
0.18M
0.12M
0.15M
0.07M
0.07M
0.10M
0.13M
0.22M
0.42M
0.13M
January
February
March
April
May
June
July
August
September
October
November
December
Month

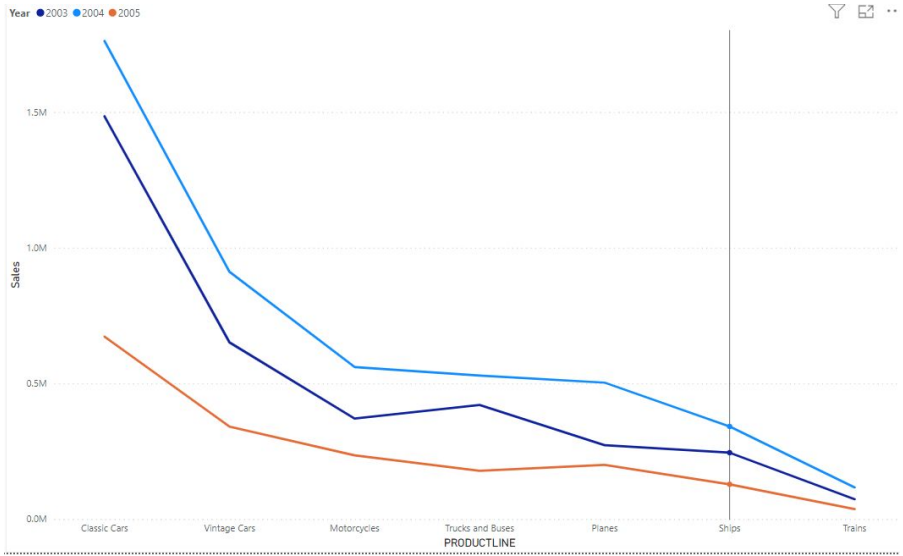

Truck and Buses
Sales
0.2M
0.1M
0.0M
0.08M
0.07M
0.06M
0.03M
0.15M
0.06M
0.06M
0.05M
0.09M
0.12M
0.25M
0.10M
January
February
March
April
May
June
July
August
September
October
November
December
Month


Motorcyycles
Sum of Motorcycles
0.2M
0.1M
0.0M
0.08M
0.12M
0.06M
0.12M
0.11M
0.05M
0.06M
0.11M
0.05M
0.10M
0.26M
0.05M
January
February
March
April
May
June
July
August
September
October
November
December
Month

Some of the common visualization techniques

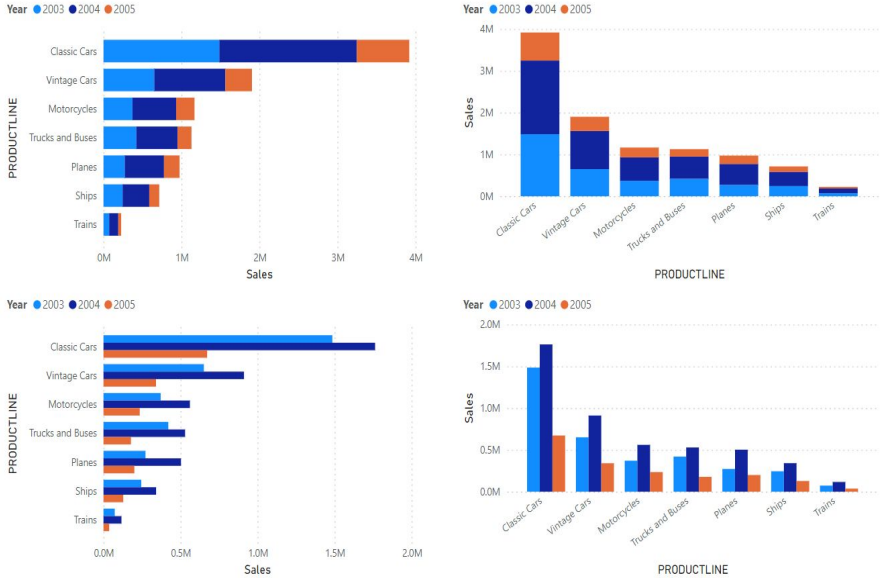
Line Plot

- Visualizing trends
- Identifying patterns/Seasonality
- Comparison
- Handling anomalies/outliners
- Forecasting and Predictive analysis
- Communicating Insights
- Tracking KPI



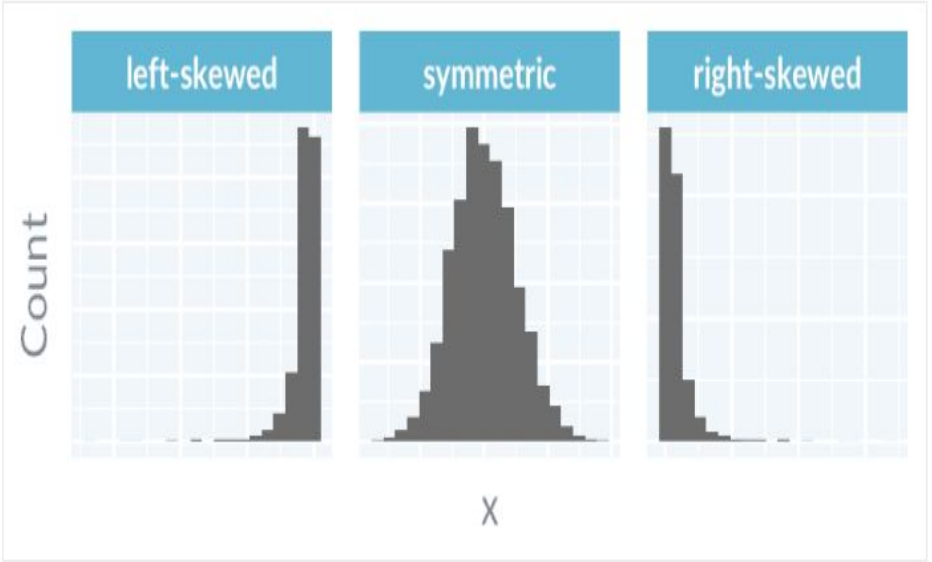
Bar Graphs

- Comparison
- Visualisation Distribution
- Highlights Trends
- Ordering/Ranking
- Forecasting and Predictive analysis
- Communicating Insights
- Tracking KPI



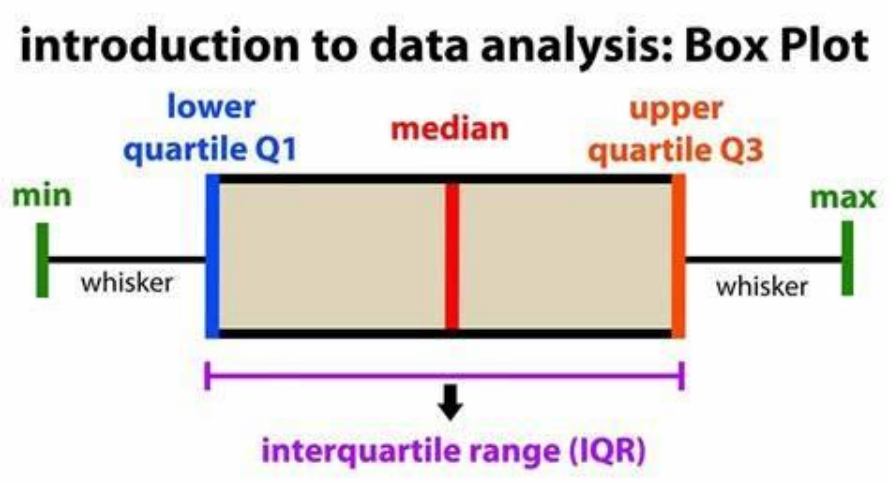
Histogram

- Visualisation Distribution·
- Identifying patterns(symmetric, skewed,unimodal, bimodal, or multimodal)
- Detecting outliers
- Assessing Data Quality
- Comparing Distributions



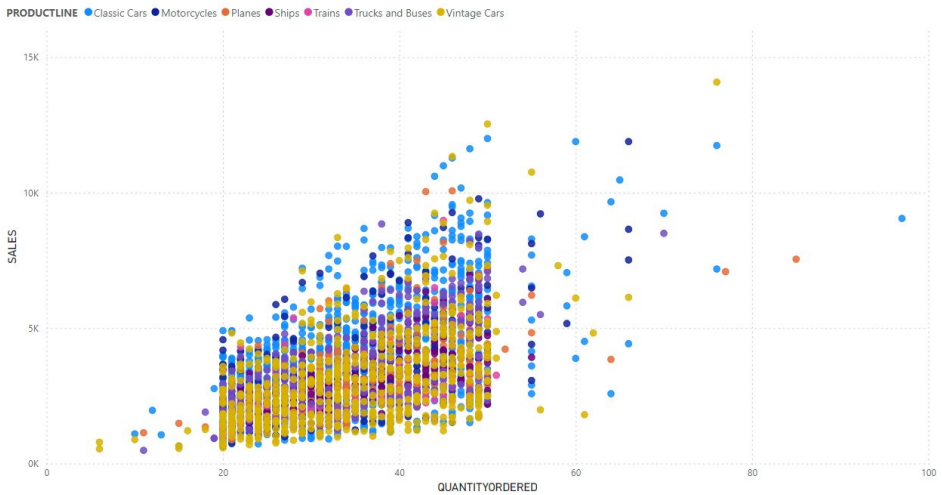
Box and Whisker plots

- Visualisation Distribution
- Identifying Outliers
- Comparison
- Assessing Symmetry and Skewness
- Variability



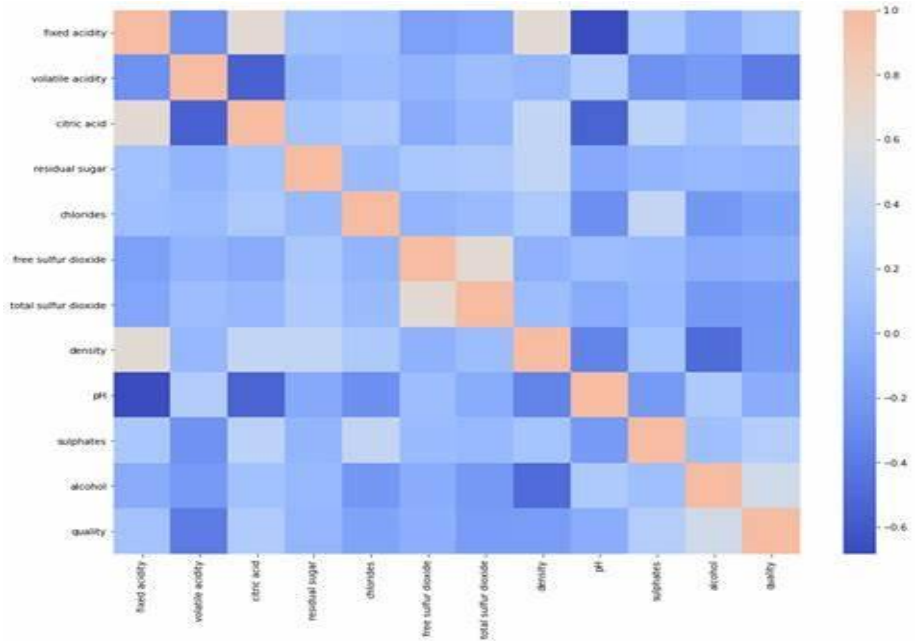
Scatter Plot

- Visualisation Distribution
- Identifying Patterns and trends
- Assessing Correlation
- Detecting Outliers
- Trend analysis
- Modeling and Prediction



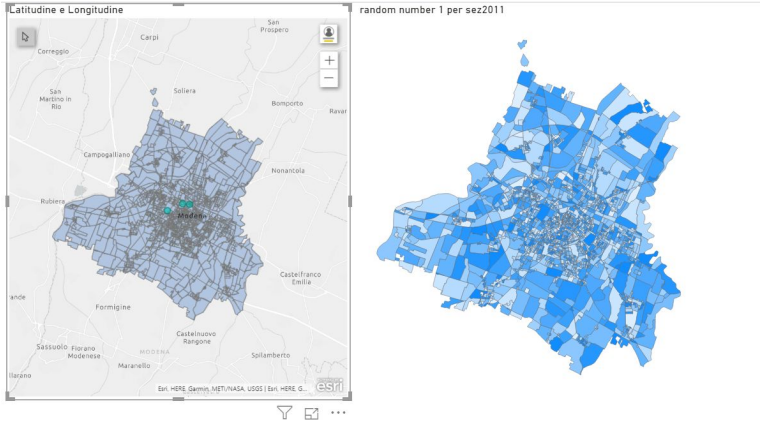
Heatmap

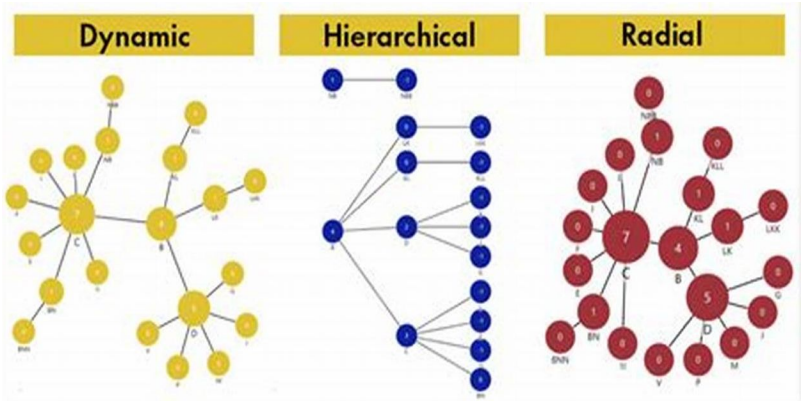
- Visualisation
- Relationships/Distributions
- Identifying Patterns
- Comparing Data
- Identifying Clusters



Other Visualization Techniques

- Maps
- Network Diagram
- Word Clouds etc





Approach to Visualization

KEY QUESTIONS?

- What **TYPE OF DATA** are you working with?
- What do you want to **COMMUNICATE**?
- Who is the **END USER** and What do you want them to?

What **TYPE OF DATA** are you working with?

- Time Series
- Geospatial
- Categorical
- Hierarchical
- Finance
- Funnel
- Survey

The type of data will define **which type of visual with best represent it**.

What do you want to **COMMUNICATE**?

Comparison

Used to compare values over time and across categories

Common Visuals

- Bar charts
- Clustered chart
- Line charts
- Area charts

Composition

Used to break down the component part of the whole

Common Visuals

- Stacked Bar/Column
- Pie/Donut chart
- Tree Maps/Sunburst
- Stacked Area charts
- Maps

Distribution

Used to show frequency of values

Common Visuals

- Histogram
- Density Plot
- Box and Whiskers
- Data table
- Heat map

Relationships

Used to show correlation between variables

Common Visuals

- Scatter plot
- Bubble chart
- Data table
- Heat map
- Correlation matrix

Keep it simple and focus on essential charts. Variation of line, bar and column charts should be sufficient for 80% of most analyses.

Who is the **END USER** and What do you want them to?

Colleagues

Detailed and granular level

- Tables or Combo charts
- Granular details to support analysis

Managers

Summarized data with clear actionable insights

- Common charts and graphs, some detail but only when it supports insights

Executives

High level crystal clear to the point information

- KPI cards
- Simple charts
- Minimal details, unless it's critical

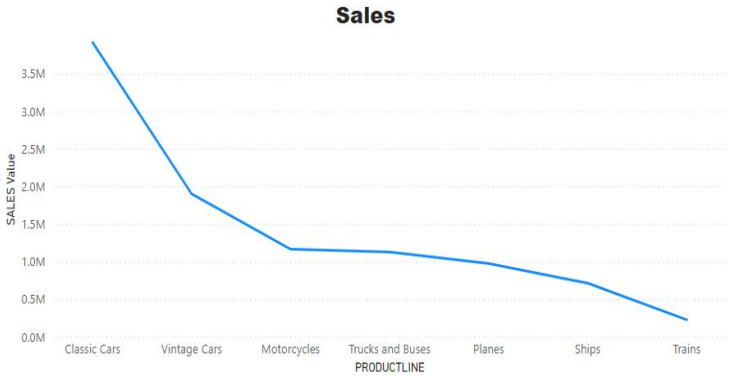
How you visualize the data is a function of **who will consume it.**

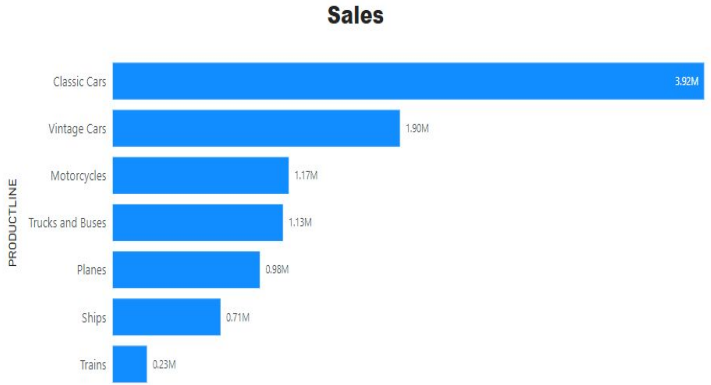
Context Is Key

Context gives meaning to data and helps user interpret them accurately.

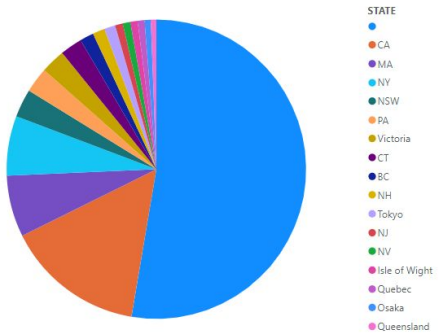
Common Data Visualization Mistakes

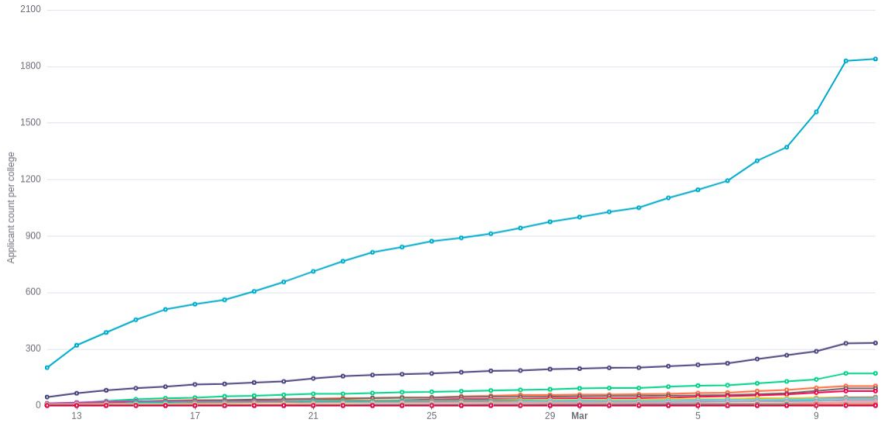
Choosing the wrong chart type



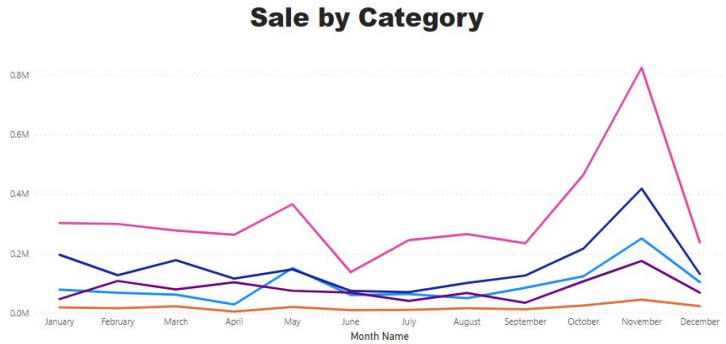


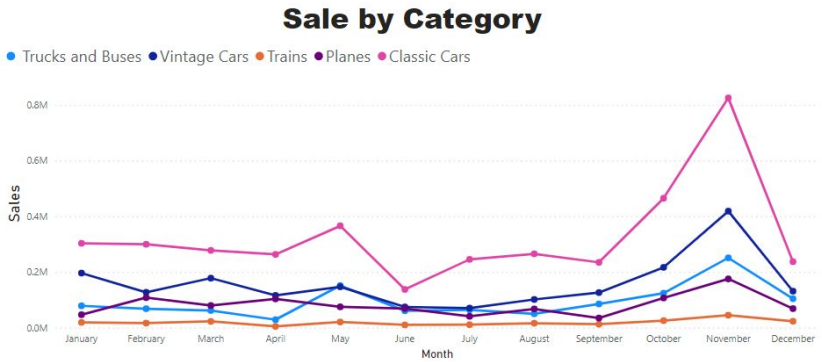
Showing too many series in single visual



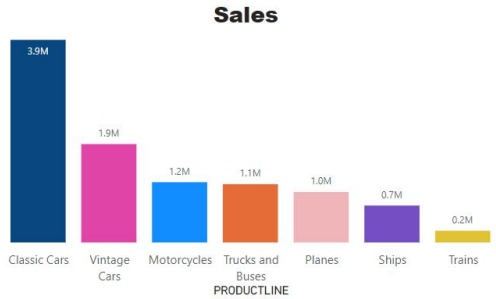


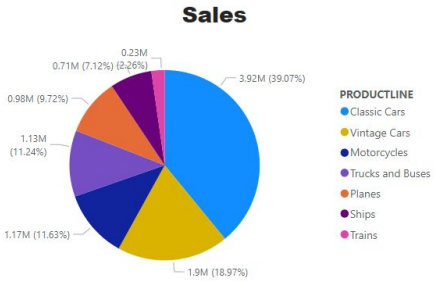
Failing to provide meaningful context

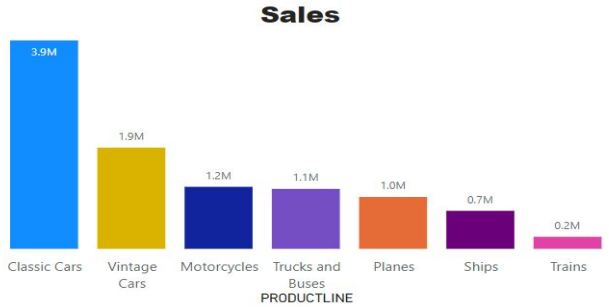




Showing too many series in single visual







Data Visualisation Practices

- Know Your Audience
- Choose the Right Chart Type
- Simplify
- Use Consistent Design
- Provide Context
- Emphasize Key Insights
- Tell a Story
- Ensure Accessibility
- Iterate and Test
- Consider Interactivity
- Optimize Performance
- Stay Up-to-Date

Seaborn

- built upon matplotlib
- Aesthetically pleasing
- simplifies visualization
- Work directly with Pandas Dataframe reducing the need for manual data manipulation

Dashboard

Dashboard are analytics tools designed to consolidate data from multiple sources track metrics and facilitate data driven decision making.

Dashboard Design

Define the purpose

- Dashboards can be made to serve various purposes like executive level report, exploratory or explanatory analysis, infographic style storytelling.
- How you approach the design of the dashboard is function of the purpose and the audience.

Key Considerations

- Who will be the end user
- What are their Key business goals and objectives
- What are the most important question they need answer to
- How frequently will the dashboard be reviewed

Choose the right metrics

- It’s important to identify which metrics and KPI to include
- Focus on metrics that directly align with level of details most appropriate for the audience

Key Considerations

- Accurate measures
- Level of details

Present data effectively

- Dashboard is all about information quickly and clearly
- Use charts and visuals best suited for the type of data

Key Considerations

- It’s all about what tell the story most effectively and efficiently
- Prioritize effectiveness over variety

Eliminate clutter

- When it comes to dashboard space is precious
- Avoid all the glitters they distract the user

Key Considerations

- Clarity over aesthetics

Use layout to focus attention

- A good layout is key to guide insights and trends upfront to guide through a logical story

Key Considerations

- Don't expect users to connect the dots on their own
- Pre attentive attributes like color, size , position etc to highlight key data
- Consider common reading pattern (Z or F patterns)

Tell a clear story

- Number, charts and graph are important but we are inspired by stories
- If your dashboard doesn't inspire action or facilitate change, What purpose does it serve.

Key Considerations

- Dashboards should be designed to tell clear data driven stories designed to expose key insights and inspire stakeholders to act.

BI tools

- Microsoft Power BI
- Tableau
- Google Looker studio
- Apache superset/preset
- Amazon quicksight

Power BI

Power BI, a business analytics service by Microsoft, furnishes interactive visualizations and business intelligence capabilities. It empowers users to interface with diverse data sources, standardize, model, and generate insightful reports and dashboards.

Power BI version

Power Bi Desktop

Resides local system and used to build and publish reports

Power Bi Web service

Resides on the web as a form of website and used to build report, make dashboards and share with others

Power Bi App

Its an app version of power BI which is used to view and share reports and dashboards

Power BI view

- Report View
- Table View
- Data model View