Chapter(3)

Non Spatial data Visualization

Lecturer: Er. Dhiraj Pyakurel

OVERVIEW

- Visualization of one, two and multi-dimensional data, Tabular data, quantitative values (scatter plot)
- Separate, Order and align (Bar, staked bar, dots and line charts),
- Tree data, Displaying Hierarchical structures,
- Graph data, Rules for graph drawing and labeling
- Time series data, Characteristics of time data, Visualization time series data, Mapping of time

ATTRIBUTE DATA VISUALIZATION

- Attribute data visualization is a way of representing data through visual elements such as graphs, charts, and diagrams.
- This type of visualization is useful for displaying data that has attributes or characteristics such as categories, time periods, and numerical values.
- Attribute data visualization is an essential tool for displaying data in a graphical format that enables easy interpretation and analysis.
- Understanding the terminology associated with attribute data visualization is essential to making meaningful interpretations of the visualizations.
- Data Types
  - Items
    - An item is an individual discrete entity
    - e.g. row in a table, node in a network
  - Attributes
    - An **attribute** is some specific property that can be measured,
    - observed, or logged
    - a.k.a. variable, (data) dimension

ATTRIBUTE DATA VISUALIZATION

- Here are some of the terminologies associated with attribute data visualization:
  - Categorical Data: This is a type of data that can be sorted into categories or groups. Examples include colors, names, and genres. Categorical data can be visualized using techniques such as bar charts and pie charts.
  - Numerical Data: This is a type of data that is represented by numbers. Examples include height, weight, and temperature. Numerical data can be visualized using techniques such as scatter plots and line graphs.
  - Time-Series Data: This is a type of data that is collected over a period of time. Examples include stock prices, weather patterns, and sales data. Time-series data can be visualized using techniques such as line graphs and heat maps.
  - Data Points: This refers to individual pieces of data in a dataset. Data points can be represented using different shapes such as dots or bars in a graph.

ATTRIBUTE DATA VISUALIZATION

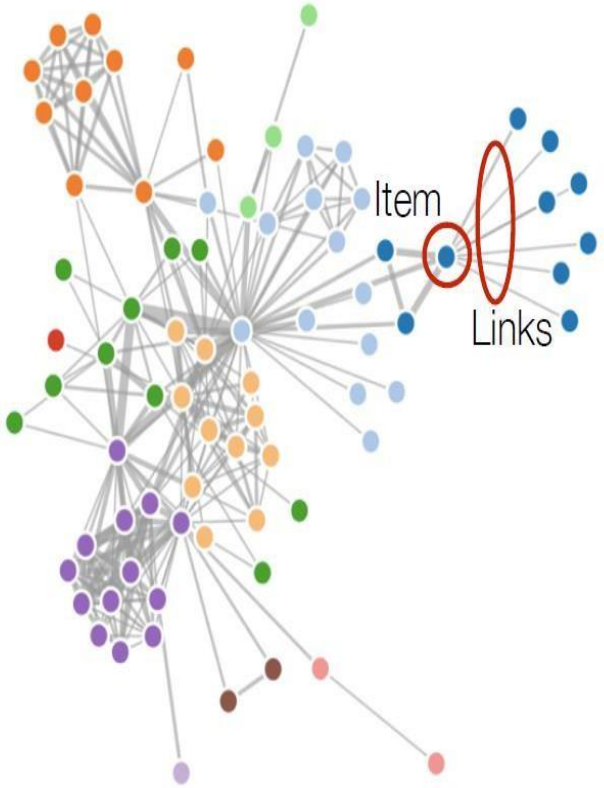
- Here are some of the terminologies associated with attribute data visualization:
  - Axis: This is the line in a graph that represents the values of one or more variables. A graph typically has two axes: a horizontal x-axis and a vertical y-axis.
  - Legend: This is a part of a graph that explains the meaning of different colors or shapes used to represent data points.
  - Trend line: This is a line that shows the general trend in a set of data. Trend lines can be added to a graph to help identify patterns and correlations.
  - Correlation: This is a measure of how strongly two variables are related. Correlation can be positive, negative, or neutral. Correlation can be visualized using techniques such as scatter plots.
  - Heat Map: This is a type of graph that represents data as colors on a grid. Heat maps are commonly used to display data that has two dimensions such as as time and geography.

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box		7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	[illegible]	5-Low	Wrap Bag	0.56	1/20/05
69	[illegible]	4-Not Specified	Small Pack	0.44	6/6/05
69	[illegible]	4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

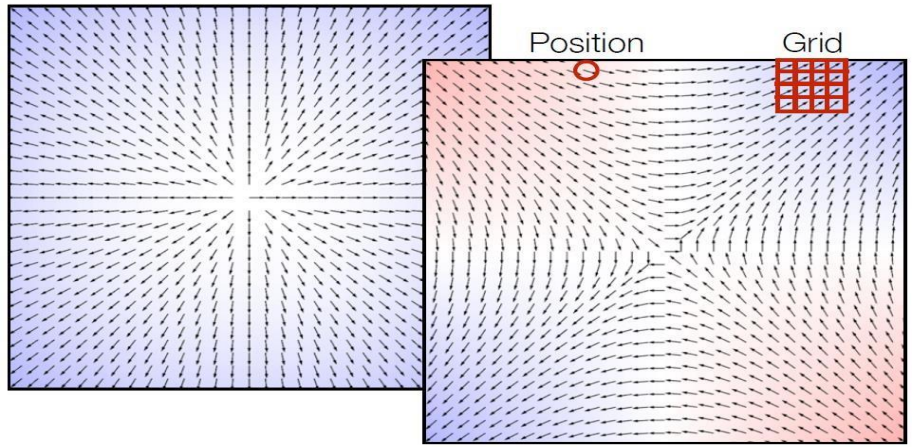
item

attribute

- Nodes
  - Synonym for item but in the context of networks (graphs)
- Links
  - A link is a relation between two items
  - e.g. social network friends, computer network links

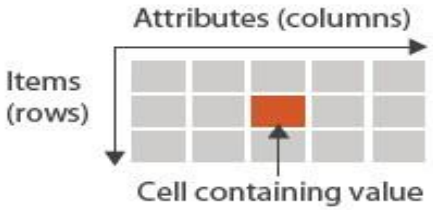


- Positions:
  - A position is a location in space (usually 2D or 3D)
  - May be subject to projections
  - e.g. cities on a map, a sampled region in an CT scan
- Grids:
  - A grid specifies how data is sampled both geometrically and topologically
  - e.g. how CT scan data is stored

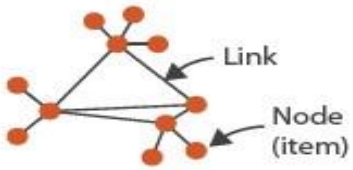


DATASET TYPES

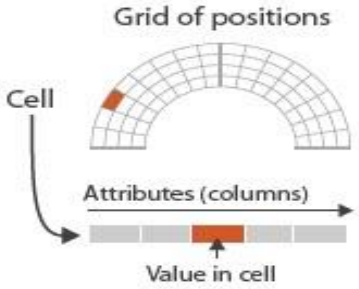
- Dataset Types
  - Tables



  - Networks

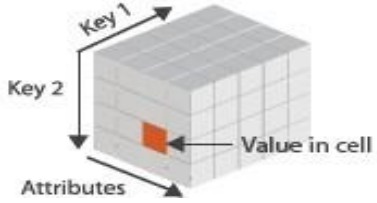


  - Fields (Continuous)



  - Geometry (Spatial)

  - *Multidimensional Table*



  - *Trees*

ONE-DIMENSIONAL DATA VISUALIZATION

- This type of data visualization involves displaying data that has one variable.
- One-dimensional data visualization is typically used for displaying categorical or numerical data.
- Examples of one-dimensional data visualization techniques include bar charts, pie charts, and histograms.
  - Bar Charts: Bar charts are used to display categorical data. The height of each bar represents the number or percentage of observations in each category. Bar charts are commonly used for comparing the frequency of different categories.
  - Pie Charts: Pie charts are used to represent categorical data as a percentage of a whole. The size of each slice is proportional to the frequency or percentage of observations in each category. Pie charts are useful for visualizing data where the total of all categories is known and the categories are mutually exclusive.
  - Histograms: Histograms are used to represent numerical data. Histograms show the distribution of a dataset by grouping the data into bins and plotting the frequency of observations in each bin. Histograms are commonly used for exploring the distribution of continuous data.

ONE-DIMENSIONAL DATA VISUALIZATION

TWO-DIMENSIONAL DATA VISUALIZATION:

- This type of data visualization involves displaying data that has two variables.
- Two-dimensional data visualization is typically used for displaying numerical data.
- Examples of two-dimensional data visualization techniques include scatter plots and line graphs.
  - Scatter Plots: Scatter plots are used to display the relationship between two variables. Each data point in the scatter plot represents a pair of values for the two variables. Scatter plots are useful for visualizing patterns and correlations in the data.
  - Line Graphs: Line graphs are used to display changes in a variable over time. Line graphs show the relationship between two variables by plotting the values of one variable on the y-axis and the time on the x-axis. Line graphs are useful for visualizing trends and changes in the data over time.

MULTI-DIMENSIONAL DATA VISUALIZATION

- This type of data visualization involves displaying data that has more than two variables.
- Multi-dimensional data visualization is typically used for displaying complex numerical data.
- Examples of multi-dimensional data visualization techniques include bubble charts and heat maps.
  - Bubble Charts: Bubble charts are used to display three-dimensional data by representing the values of three variables as the x-axis, y-axis, and size of the bubble. Bubble charts are useful for visualizing complex relationships between three variables.
  - Heat Maps: Heat maps are used to display data as a two-dimensional grid of colored cells. The color of each cell represents the value of the variable for that cell. Heat maps are useful for visualizing large datasets where patterns and trends are not immediately apparent.

TABULAR DATA

- Tabular data is data that is organized in rows and columns in a table format.
- Tabular data is often used to represent quantitative values, such as numerical data or percentages.
- Scatter plots are a type of data visualization that can be used to represent quantitative values.
- Scatter plots display the relationship between two variables, where one variable is plotted on the x-axis and the other variable is plotted on the y-axis.
- Each point on the scatter plot represents a pair of values for the two variables.
- In data visualization, tabular data can be used to create charts, graphs, and other visualizations that help users to better understand the data.
- For example, let's say you have a table of sales data for a company, organized by region, product, and time period. You could use this data to create a line chart showing sales trends over time for each region, or a bar chart showing total sales by product.

TABULAR DATA

```
import pandas as pd
import matplotlib.pyplot as plt

# Define the data as a Pandas DataFrame
data = pd.DataFrame({
    'Product': ['Product A', 'Product B', 'Product C', 'Product D'],
    'Sales': [100, 150, 75, 200]
})

# Create a bar chart using the data
plt.bar(data['Product'], data['Sales'])

# Add labels and a title
plt.xlabel('Product')
plt.ylabel('Sales')
plt.title('Total Sales by Product')

# Display the chart
plt.show()
```

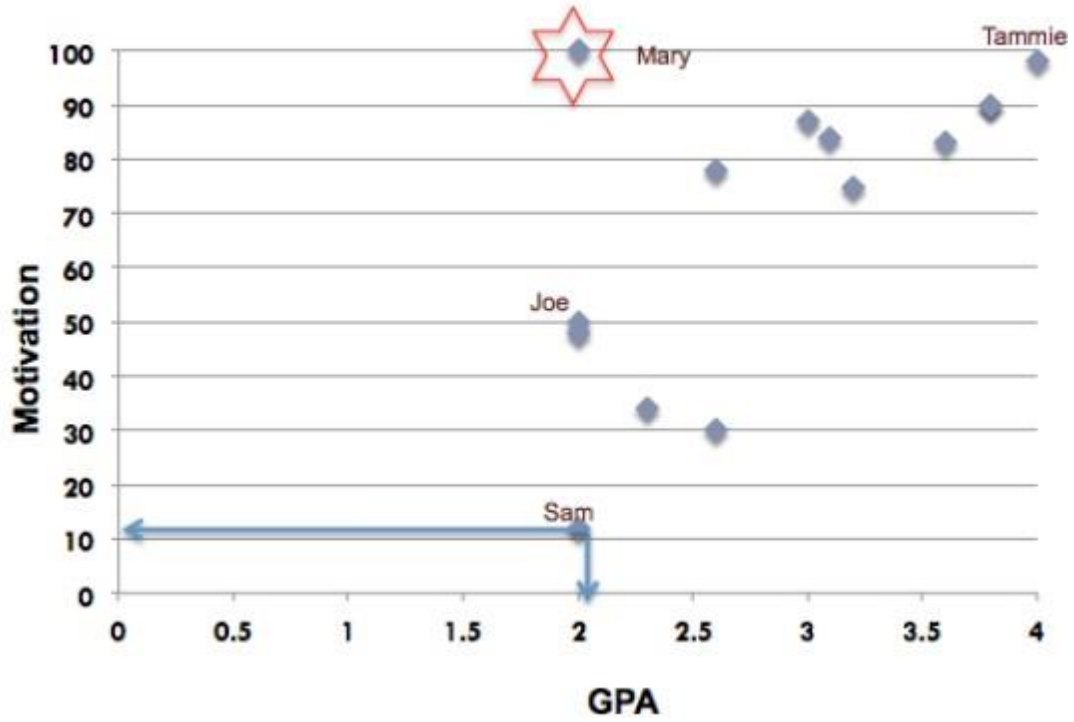
QUANTITATIVE VALUES (SCATTER PLOT)

- Quantitative values are numerical values that can be measured and analyzed mathematically.
- In data visualization, scatter plots are a useful way to visualize relationships between two quantitative values.
- A scatter plot is a graph in which the values of two quantitative variables are plotted along two axes, and each data point is represented by a dot on the graph.
- The position of each dot on the graph indicates the values of the two variables for that data point, and patterns or trends in the data can often be seen by examining the overall shape of the scatter plot.
- Scatter plots can be very useful for exploring relationships between two quantitative variables.
- In addition to examining overall patterns in the data, you can also use techniques like regression analysis to identify more specific relationships between the variables.

QUANTITATIVE VALUES (SCATTER PLOT)

What is the relationship between students' achievement motivation and GPA?

Student	Student GPA	Motivation
Joe	2.0	50
Lisa	2.0	48
Mary	2.0	100
Sam	2.0	12
Deana	2.3	34
Sarah	2.6	30
Jennifer	2.6	78
Gregory	3.0	87
Thomas	3.1	84
Cindy	3.2	75
Martha	3.6	83
Steve	3.8	90
Jamell	3.8	90
Tammie	4.0	98



QUANTITATIVE VALUES (SCATTER PLOT)

```
import pandas as pd
import matplotlib.pyplot as plt

# Define the data as a Pandas DataFrame
data = pd.DataFrame({
    'Age': [25, 31, 38, 42, 57, 61, 65, 71],
    'Income': [34000, 44000, 52000, 61000, 74000, 88000, 98000, 106000]
})

# Create a scatter plot using the data
plt.scatter(data['Age'], data['Income'])

# Add labels and a title
plt.xlabel('Age')
plt.ylabel('Income')
plt.title('Income by Age')

# Display the chart
plt.show()
```

- In this example, we have a set of data that includes the ages and incomes of several individuals.
- We use Pandas to define this data as a DataFrame with one column for age and one column for income.
- We then use Matplotlib to create a scatter plot of this data, with age on the x-axis and income on the y-axis.
- The resulting chart shows that there is a general trend of increasing income with increasing age, although there is also some variability in the data.

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

- Separate, order, and align are techniques used in data visualization to improve the clarity and effectiveness of charts and graphs

1. Separate: This technique involves separating different groups of data to make it easier to compare and contrast them. For example, you might use separate bars or colors to represent different categories of data. This can be especially useful when comparing data across different time periods or locations.
   - Here's an example of a bar chart that separates data by category:
   - In this example, we use a bar chart to compare sales data for four different categories (A, B, C, and D) in two different years (2020 and 2021). We use different colors for the bars representing each year to make it easier to compare the data.

```
import pandas as pd
import matplotlib.pyplot as plt

# Define the data as a Pandas DataFrame
data = pd.DataFrame({
    'Category': ['A', 'B', 'C', 'D'],
    'Sales 2020': [100, 150, 75, 200],
    'Sales 2021': [125, 175, 90, 220]
})

# Create a bar chart using the data
data.plot(x='Category', kind='bar')

# Add labels and a title
plt.xlabel('Category')
plt.ylabel('Sales')
plt.title('Sales by Category')

# Display the chart
plt.show()
```

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

- Order: This technique involves ordering data along an axis to highlight patterns or trends.
- For example, you might order data by value or by date to show changes over time.
- Here's an example of a stacked bar chart that orders data by value:
- In this example, we use a stacked bar chart to compare sales data for four different regions (East, West, North, and South) in four different quarters of the year. We order the data by Q4 sales to highlight which regions had the highest sales.

```
import pandas as pd
import matplotlib.pyplot as plt

# Define the data as a Pandas DataFrame
data = pd.DataFrame({
    'Region': ['East', 'West', 'North', 'South'],
    'Q1': [100, 150, 75, 200],
    'Q2': [125, 175, 90, 220],
    'Q3': [150, 200, 110, 240],
    'Q4': [175, 225, 125, 260]
})

# Order the data by Q4 sales
data = data.sort_values(by='Q4')

# Create a stacked bar chart using the data
data.plot(x='Region', kind='bar', stacked=True)

# Add labels and a title
plt.xlabel('Region')
plt.ylabel('Sales')
plt.title('Sales by Region and Quarter')

# Display the chart
plt.show()
```

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

- Align: This technique involves aligning different data points along a common axis to make it easier to compare them. For example, you might align data points by date or by a specific value to show changes over time or differences between categories.
- Here's an example of a line chart that aligns data points by date:

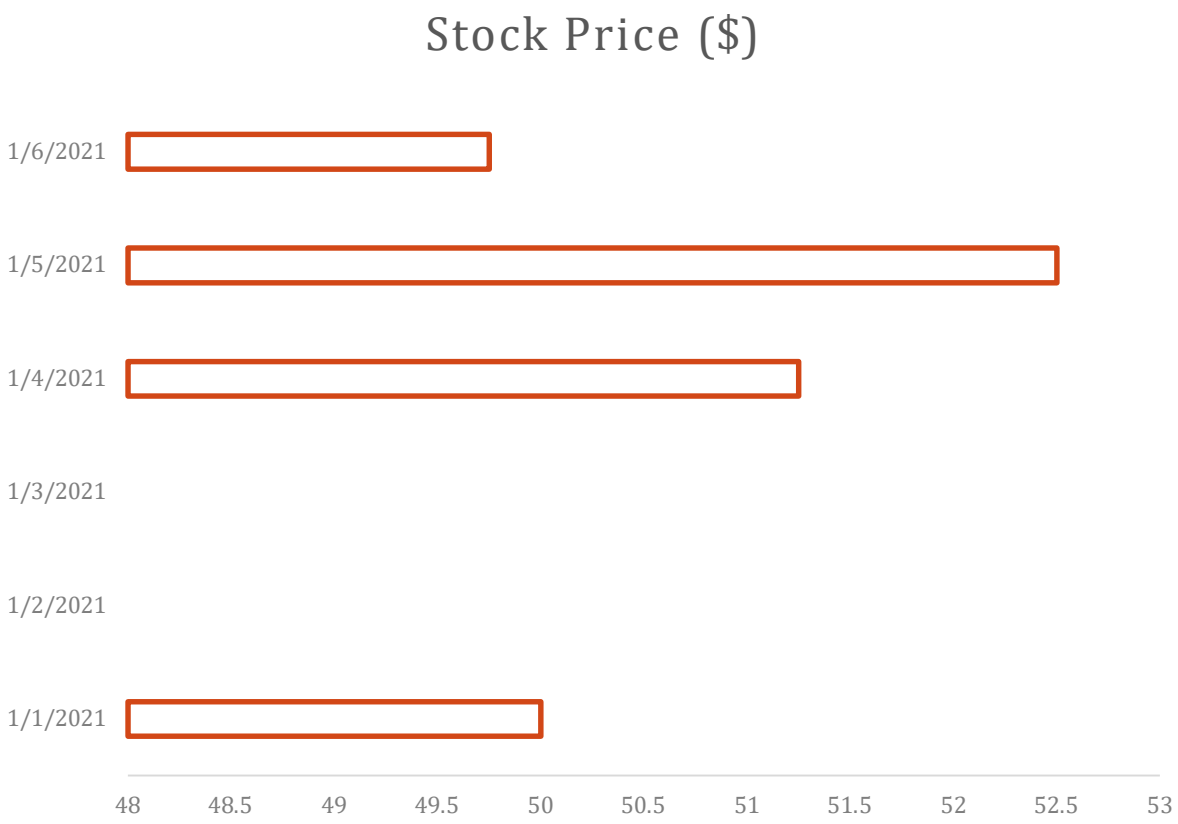
```
import pandas as pd

import matplotlib.pyplot as plt

# Define the data as a Pandas DataFrame

data = pd.DataFrame({ 'Date': ['2020-01-01', '2020-02-01', '2020-03-01', '2020-04-01', '2020-05-01'],

   'Sales': [100, 125, 150, 175, 200],

   'Expenses':[75, 100, 125, 150, 175] })

# Convert the Date column to a datetime object

data['Date'] = pd.to_datetime(data['Date'])

# Create a line chart using the data

data.plot(x='Date', kind='line')

# Add labels and a title

plt.xlabel('Date') plt.ylabel('Amount') plt.title('Sales and Expenses by Month')

# Display the chart

plt.show()
```

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

Bar charts:

- Bar charts are a common type of chart used to display numerical data using horizontal or vertical bars. They are effective for comparing discrete values across different categories or groups.
- Example: A vertical bar chart can be used to compare the number of students enrolled in different courses. By separating the bars by color and ordering them by enrollment size, viewers can quickly identify which courses are most popular and which ones are less popular.

BAR CHART



Course	Number of students enrolled
Math	120
English	90
Science	80
History	60

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

- Stacked bar charts:
- Stacked bar charts are similar to bar charts, but they are used to show how different components contribute to a whole. Each bar is divided into segments that represent different categories or sub-groups, with each segment representing a percentage of the whole.
- Example: A stacked bar chart can be used to show how different age groups contribute to the overall population of a city. By separating the bars into age groups and ordering them by age, viewers can quickly see how the population is distributed across different age groups.

STAKED CHARTS

Age Group	Population
0-9	50000
10-19	75000
20-29	100000
30-39	90000
40-49	80000
50-59	70000
60-69	50000
70+	40000

In this chart, each bar would be separated into eight segments, one for each age group, and each segment would be labeled with the age range it represents. The bars would be ordered by age group, with the youngest age group (0-9) at the bottom and the oldest age group (70+) at the top.
By using a stacked bar chart in this way, viewers can quickly see how the population of a city is distributed across different age groups and how each age group contributes to the overall population. This type of chart can be useful for demographic analysis and planning purposes

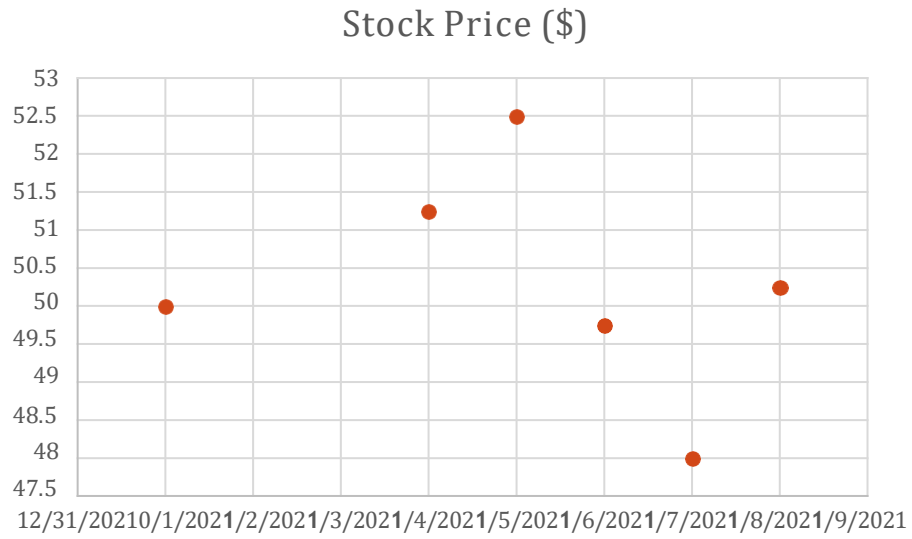
SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

Dot plots:

- Dot plots are used to display numerical data using a series of dots aligned along a common axis. They are effective for displaying small datasets and identifying outliers or anomalies.
- Example: A dot plot can be used to compare the heights of different individuals. By aligning the dots along a common baseline and ordering them by height, viewers can quickly see how the heights of different individuals compare to one another.

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

Individual	Height (inches)
Alice	65
Bob	72
Carol	68
Dave	66
Eve	64
Frank	69



In this chart, each dot would represent an individual, and the dots would be ordered by height, with the tallest individual (Bob) at the top and the shortest individual (Eve) at the bottom. By using a dot plot in this way, viewers can quickly see how the heights of different individuals compare to one another. This type of chart can be useful for visualizing and analyzing various types of data, including height, weight, and other quantitative measures

SEPARATE, ORDER AND ALIGN (BAR, STAKED BAR, DOTS AND LINE CHARTS),

Line charts:

- Line charts are used to display trends in numerical data over time. They are effective for identifying patterns and trends in the data, as well as highlighting changes over time.
- Example: A line chart can be used to track the performance of a company's stock over time. By aligning the data points along a common axis and ordering them by date, viewers can quickly see how the stock price has fluctuated over time and identify any significant changes or trends.

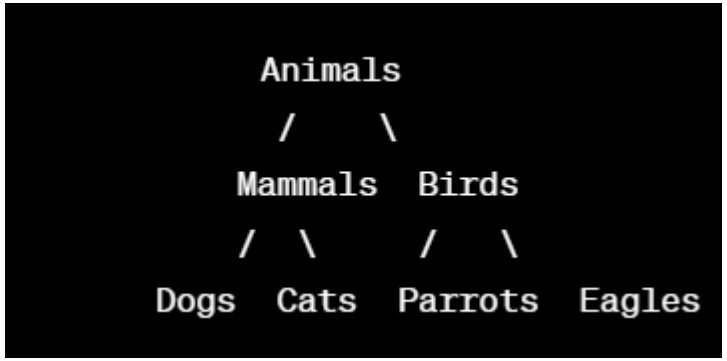
TREE DATA, DISPLAYING HIERARCHICAL STRUCTURES

Displaying hierarchical structures in data visualization is important for understanding relationships and patterns in data. One common way to display hierarchical structures is through tree diagrams or tree maps.

- Tree data is a common type of data that represents hierarchical relationships between objects.
- Tree data is used to represent a wide range of information, from organizational structures to computer file systems.
- The nodes of a tree data structure represent the objects in the hierarchy, while the edges represent the relationships between them.
- In data visualization, displaying hierarchical structures is commonly done using a tree diagram.
- A tree diagram is a graphical representation of a tree data structure that shows the relationships between nodes. It typically consists of nodes, edges, and labels
- Nodes in a tree diagram represent objects or groups of objects, while edges represent the relationships between them.
- Each node can have one or more child nodes, representing the objects or groups of objects that are "contained" within it. The root node represents the top-level object or group of objects in the hierarchy.

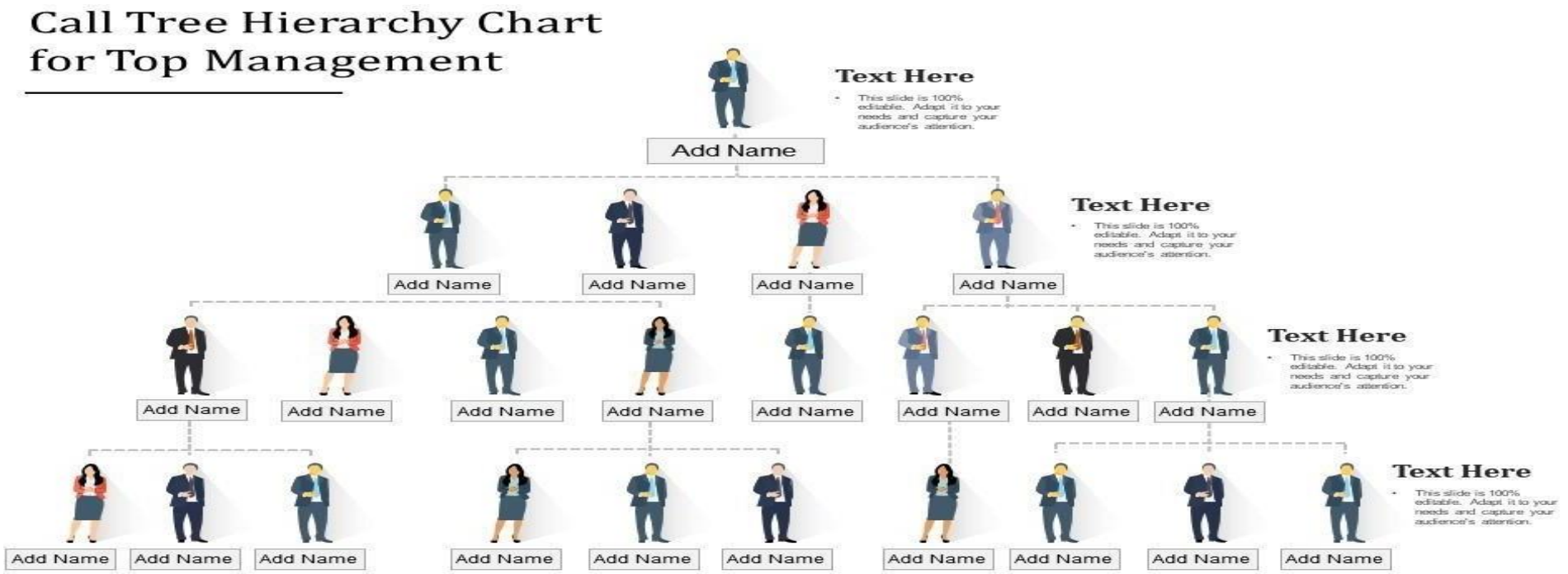
- For example, a company's organizational structure can be displayed using a tree diagram. The top-level node represents the CEO or president, and the child nodes represent different departments or divisions. Each department can have further child nodes representing teams or individuals

TREE DATA, DISPLAYING HIERARCHICAL STRUCTURES



- In this example, the "Animals" node is the root node, representing the top-level object in the hierarchy. It has two child nodes, "Mammals" and "Birds", which represent groups of objects that are "contained" within the "Animals" node.
- The "Mammals" node has two child nodes, "Dogs" and "Cats", which represent groups of objects that are "contained" within the "Mammals" node. Similarly, the "Birds" node has two child nodes, "Parrots" and "Eagles".
- Each node in the tree diagram can be labeled with additional information about the objects it represents. For example, the "Dogs" node might be labeled with the number of dogs in the group, or the breeds of dogs in the group.

TREE DATA, DISPLAYING HIERARCHICAL STRUCTURES



GRAPH DATA

- Graph data is a type of data that represents relationships between objects.
- A graph consists of nodes, also known as vertices, and edges, which connect the nodes.
- In data visualization, graph data is often displayed using graph diagrams, also known as network diagrams.
- Graphs are powerful tools for data visualization because they allow us to see patterns and relationships in data that may not be immediately obvious when looking at the raw numbers
- There are several rules for graph drawing that can help create clear and informative visualizations:
  - Avoid crossing edges: Crossing edges can make a graph difficult to read and understand. It's best to try to arrange the nodes and edges in a way that minimizes edge crossings.
  - Use symmetry and alignment: Symmetry and alignment can help create a sense of order in a graph. For example, arranging nodes in a circular or radial pattern can help create symmetry, while aligning nodes or edges can help create a sense of order.
  - Use color and size to distinguish nodes: Using color and size can help distinguish different types of nodes in a graph. For example, nodes representing different categories can be color-coded, while larger nodes can represent nodes with more connections.
  - Label nodes and edges: Labeling nodes and edges can provide additional information about the objects and relationships in a graph. Labels can be used to display names, values, or any other relevant information.

RULES FOR GRAPH DRAWING AND LABELING

- Here are some examples of how to follow the rules for graph drawing and labeling in data visualization:
  - Choose the appropriate graph type: Suppose you want to visualize the sales performance of a company over time. In this case, a line graph would be appropriate as it allows you to show trends in sales over a period of time.
  - Use clear and concise labels: When labeling the axes, use clear and concise labels such as "Year" for the x-axis and "Sales" for the y-axis. Use a clear and descriptive title such as "Sales Performance of XYZ Company (2015-2022)".
  - Scale the axes appropriately: Ensure that the scale on the y-axis is appropriate for your data. For example, if the sales figures range from $10,000 to $100,000, set the y-axis scale to show all values in this range.
  - Avoid clutter: Avoid cluttering the graph with unnecessary information such as gridlines or multiple legends. Use only the essential elements that help to convey your message.
  - Use color and visual aids carefully: Use color and visual aids to highlight important information, such as using a different color for each product line. However, avoid using too many colors or visual aids, as this can make the graph confusing.
  - Provide context: Provide context for the data by including a brief explanation of what the graph is showing and why it is important. For example, you might explain that the graph shows the sales performance of XYZ Company over the last 7 years and that it is important to understand the trends in order to make informed business decisions

TIME SERIES DATA

- Time series data refers to a type of data that is collected over a period of time at regular intervals.
- Time series data can be visualized using various types of graphs such as line charts, area charts, and candlestick charts.
- Time series data provides significant value to organizations because it enables them to analyze important real-time and historical metrics.
- A time-series describes a sequence of measurements over time. The time between those measurements is often constant. Compared to a timeline it contains a much more specific dataset and the changes in this dataset over time have to be determines and visualized.

TIME SERIES DATA

Components of Time Series:-

The change which are being in time series, They are effected by Economic, Social, Natural, Industrial & Political Reasons. These reasons are called components of Time Series.

- SECULAR TREND :-
- SEASONAL VARIATION :-
- CYCLICAL VARIATION :-
- IRREGULAR VARIATION :-

TIME SERIES DATA

❑Secular trend:

The increase or decrease in the movements of a time series is called Secular trend.

A time series data may show upward trend or downward trend for a period of years and this may be due to factors like:

- ❑ increase in population,
- ❑change in technological progress ,
- ❑large scale shift in consumers demands,

For example,

- population increases over a period of time,price increases over a period of years,production of goods on the capital market of the country increases over a period of years.These are the examples of upward trend.
- The sales of a commodity may decrease over a period of time because of better products coming to the market.This is an example of declining trend or downward.

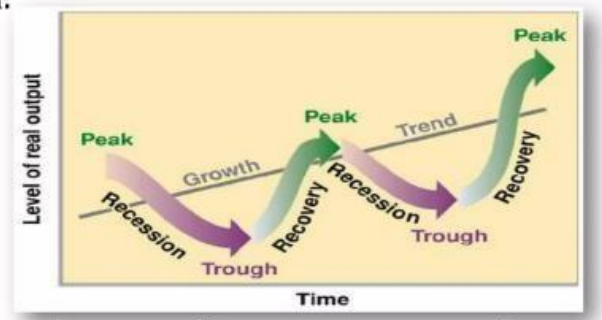
• Seasonal variation:

- Seasonal variation are short-term fluctuation in a time series which occur periodically in a year. This continues to repeat year after year.
  - The major factors that are weather conditions and customs of people.
  - More woolen clothes are sold in winter than in the season of summer .
  - each year more ice creams are sold in summer and very little in Winter season.
  - The sales in the departmental stores are more during festive seasons that in the normal days.

TIME SERIES DATA

❑Cyclical Variations:

Cyclical variations are recurrent upward or downward movements in a time series but the period of cycle is greater than a year. Also these variations are not regular as seasonal variation.



A business cycle showing these oscillatory movements has to pass through four phases-prosperity, recession, depression and recovery. In a business, these four phases are completed by passing one to another in this order.

• Irregular variation:

Irregular variations are fluctuations in time series that are short in duration, erratic in nature and follow no regularity in the occurrence pattern. These variations are also referred to as residual variations since by definition they represent what is left out in a time series after trend ,cyclical and seasonal variations. Irregular fluctuations results due to the occurrence of unforeseen events like :

- FLOODS,
- EARTHQUAKES,
- WARS,
- FAMINES

CHARACTERISTICS OF TIME

- Time has a special character to consider. It follows some rules which do not necessarily apply to other data sets.
  1. Time Is Involuntary Unlike the three space dimensions, where it is possible to just stand in one place for a while, it is not possible to stand at one point of time for a while. There is no possibility to stop time.
  2. Time Is Irreversible Time is a sequence of events, happening one after another. The future is influenced by the past, but not the other way around.What has happened can not be undone.
  3. Time Is Required Nothing can happen without time. Time has to progress in order for anything to happen.
  4. Time Is Measurable There are different units to measure time, such as seconds, minutes, hours, years, ... Important is only that it is possible to measure time, as known from the watches and clocks.
  5. Time Is Absolute Einstein told us that time is relative and strongly depends on the speed of movement. When looking at time on the earth, however, it can be said that time is absolute, which means it is the same whether it is measured in Europe or in America, at the South Pole or at the North Pole.

CHARACTERISTICS OF TIME DATA

- Characteristics of time data:
  1. Time series data is typically collected at regular intervals, such as hourly, daily, weekly, or monthly.
  2. Time series data is often subject to seasonality or trends, meaning that the data may exhibit regular patterns or fluctuations over time.
  3. Time series data may have missing values due to incomplete data collection or data errors.
  4. Time series data may exhibit autocorrelation, meaning that the values of the data are dependent on previous values.

VISUALIZATION TIME SERIES DATA

- Visualizing time series data is important to identify patterns and trends over time, detect anomalies, and make informed decisions based on the data.

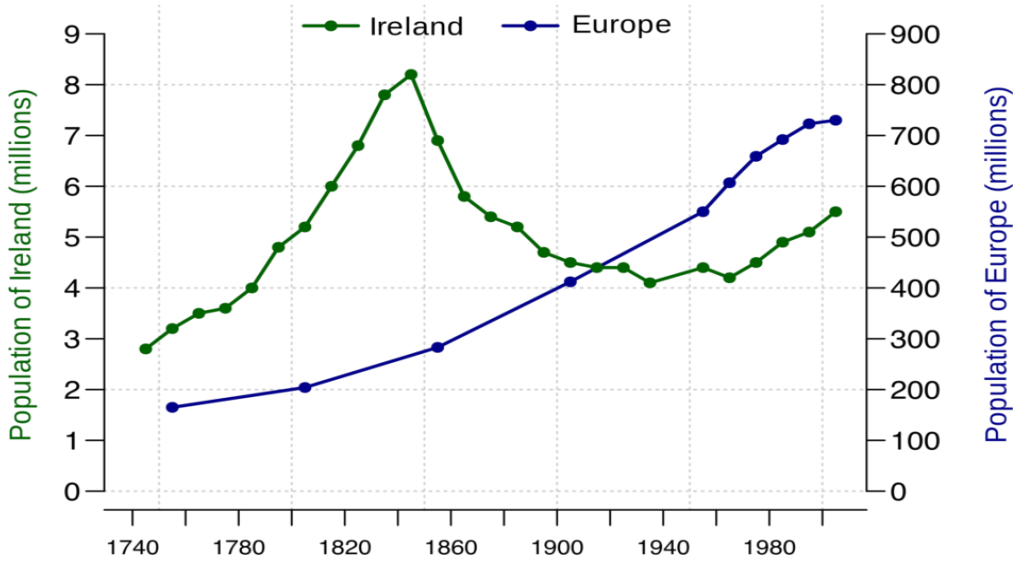
1. **Line Graph**

   A line graph uses points connected by lines (also called trend lines) to show how a dependent variable and independent variable changed.

   An independent variable, true to its name, remains unaffected by other parameters, whereas the dependent variable depends on how the independent variable changes.

   For temporal visualizations, time is always the independent variable, which is plotted on the horizontal axis. Then the dependent variable is plotted on the vertical axis.

In the graph below, the populations of Europe and Ireland are the dependent variables and time is the independent variable.



VISUALIZATION TIME SERIES DATA

- Assignment find other types with examples

MAPPING OF TIME.

- Mapping time refers to the process of converting time data into a format that can be used for analysis and visualization.
- This may involve converting the data into a standardized format, such as a timestamp or a datetime object, and accounting for missing values or outliers in the data.
- Mapping time may also involve aggregating the data into larger time intervals, such as hourly or daily averages, to make it easier to visualize and analyze trends in the data over longer periods of time

MAPPING OF TIME

- There are several ways to map time, including:

1. X-axis: In many types of time series visualizations, time is mapped to the x-axis of the chart. This allows the viewer to easily see the trend of the data over time. The x-axis can be labeled with the time period, such as year, month, or day, depending on the level of granularity of the data.
2. Color: Color can be used to represent time in data visualizations. For example, a heat map can use a color gradient to show changes in temperature over time. In this case, the color represents the value of the data point at a particular time.
3. Size: Size can also be used to represent time in data visualizations. For example, a scatter plot can use the size of the data point to represent the value of the variable at a particular time. This can help highlight outliers or trends in the data over time.
4. Animation: Animation is a powerful tool for mapping time in data visualization. By animating a chart or graph, the viewer can see how the data changes over time. For example, an animated line chart can show the trend of the data over a period of time, with the line moving along the x-axis as the time changes.
5. Interactive controls: Interactive controls, such as sliders or drop-down menus, can be used to map time in data visualizations. These controls allow the viewer to adjust the time period shown in the chart, making it easier to explore the data at different levels of granularity.

ASSIGNMENT

1. What are the advantages and disadvantages of using scatter plots to visualize quantitative values in one-dimensional data? How can these limitations be overcome?
2. Compare and contrast different types of bar charts (e.g., horizontal, vertical, stacked, grouped) and discuss the circumstances under which each type is most appropriate. Provide examples of real-world applications where each type has been successfully used to display tabular data.
3. What are the key principles that guide the visualization of hierarchical structures using tree diagrams? Discuss the different types of tree diagrams (e.g., radial, rectangular, sunburst) and evaluate their strengths and weaknesses in terms of their ability to convey the structure and relationships within the data.
4. Discuss the key principles that govern the drawing and labeling of graphs in data visualization. How can these principles be used to create effective visualizations that accurately represent the relationships between variables?

ASSIGNMENT

5. What are the characteristics of time series data, and how do these impact the choice of visualization techniques? Compare and contrast different methods for visualizing time series data, including line charts, area charts, and stacked bar charts.
6. How can data visualization be used to reveal patterns and insights in multi-dimensional data? Discuss the use of techniques such as scatter plots, heat maps, and parallel coordinate plots to identify relationships between variables and uncover hidden trends in complex datasets.
7. What are the key considerations in designing effective data visualizations for non-expert audiences? How can visualizations be tailored to different audiences and contexts to ensure that they are accessible, informative, and engaging?
8. Discuss the challenges and opportunities associated with the use of data visualization in exploratory data analysis. How can visualization be used to help analysts identify interesting patterns and trends in large datasets, and what are the limitations of these techniques?