Assignment. No.2

Semester	T.E. Semester VI – Computer Engineering
Subject	Data Warehousing and Mining
Subject Professor In-charge	Prof. Kavita Shirsat
Assisting Teachers	Prof. Kavita Shirsat

Student Name	Deep Salunkhe	
Roll Number	21102A0014	
Grade and Subject Teacher's Signature		

Assignment Number	01
Assignment Title	Visualization Techniques.

Data Visualization

Data visualization is the graphical representation of data and information to help users understand patterns, trends, and insights hidden in the data. It enables data analysts and decision-makers to communicate complex information effectively and make data-driven decisions.

**Types of Data Visualization**
- Pixel Based
- Icon Based

Major Differences Between Pixel-Based and Icon-Based Data Visualization:
1. Representation:
   - Pixel-Based: Pixel-based visualizations use graphical elements like dots, lines, and bars to represent individual data points and patterns in the data.
   - Icon-Based: Icon-based visualizations use visual symbols or icons, such as pie slices or gauge needles, to represent data categories or single values.
2. Use Cases:
   - Pixel-Based: Pixel-based visualizations are more versatile and can be used to display a wide range of data types, from continuous data in scatter plots to categorical data in bar charts.
   - Icon-Based: Icon-based visualizations are often used to represent proportions, percentages, or single values, making them suitable for showing parts of a whole or progress towards a goal.
3. Flexibility:
   - Pixel-Based: Pixel-based visualizations offer more flexibility in displaying complex data relationships, allowing analysts to customize the visualization to suit their specific needs.
   - Icon-Based: Icon-based visualizations are more straightforward and have limited customization options, making them easier to interpret and suitable for quick data representation.

Assignment. No.2

Pixel-Based Data Visualization:

1. Scatter Plots: A scatter plot is a type of pixel-based visualization that represents individual data points as dots on a two-dimensional coordinate system. Each dot's position corresponds to the values of two variables, allowing analysts to observe the relationship and distribution between them.
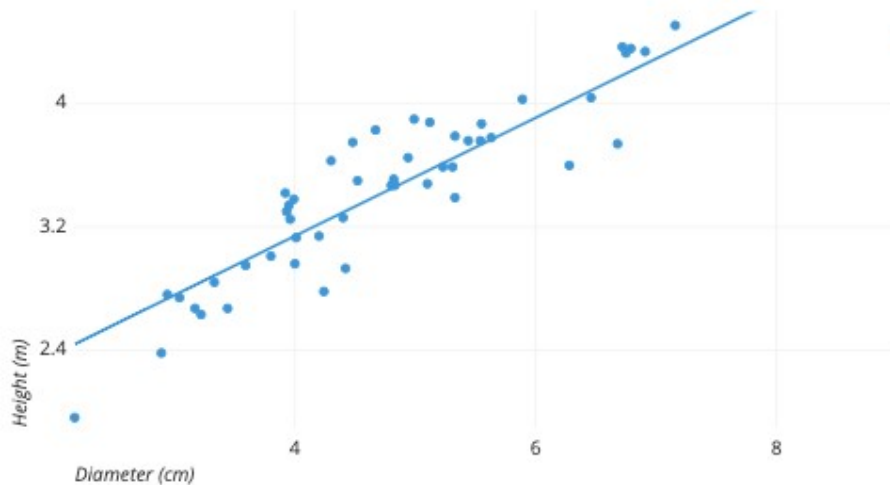

2. Line Charts: Line charts use connected data points to show trends and changes over time or continuous data. They are suitable for visualizing data with a sequence, such as stock prices or temperature variations.
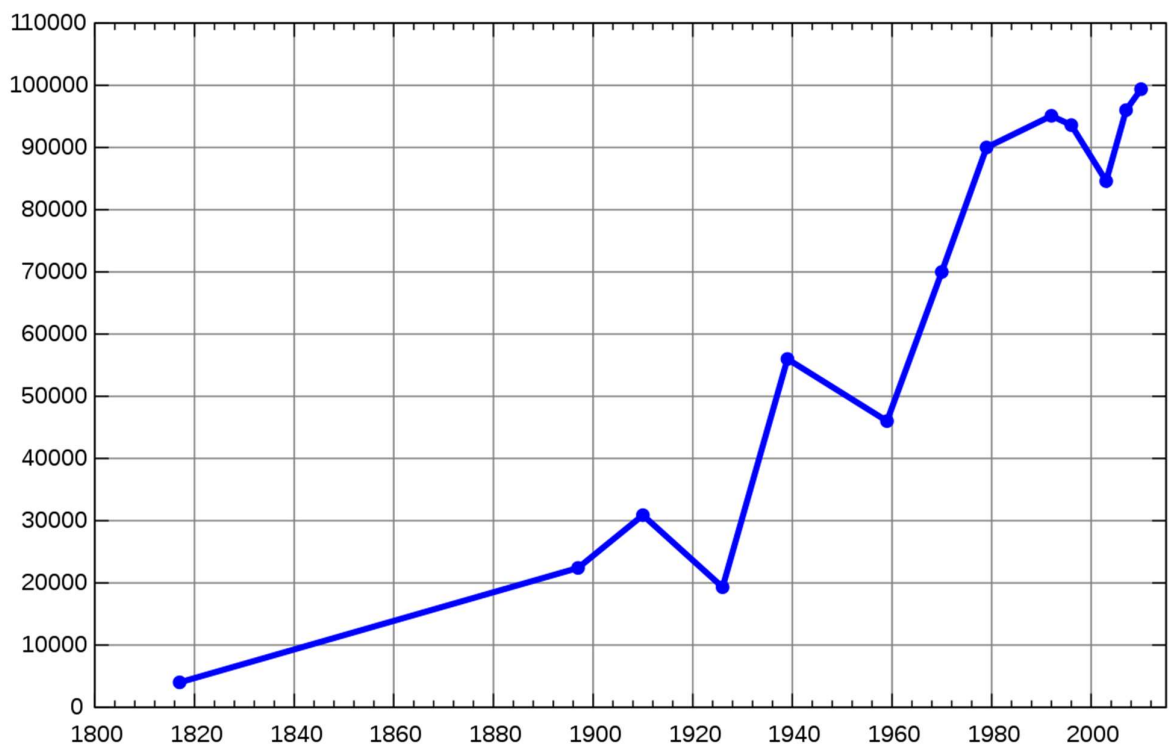
3. Bar Charts: Bar charts use rectangular bars to represent different categories or groups, and the length of each bar corresponds to the value of the data. They are effective for comparing discrete data points and can be either horizontal or vertical.
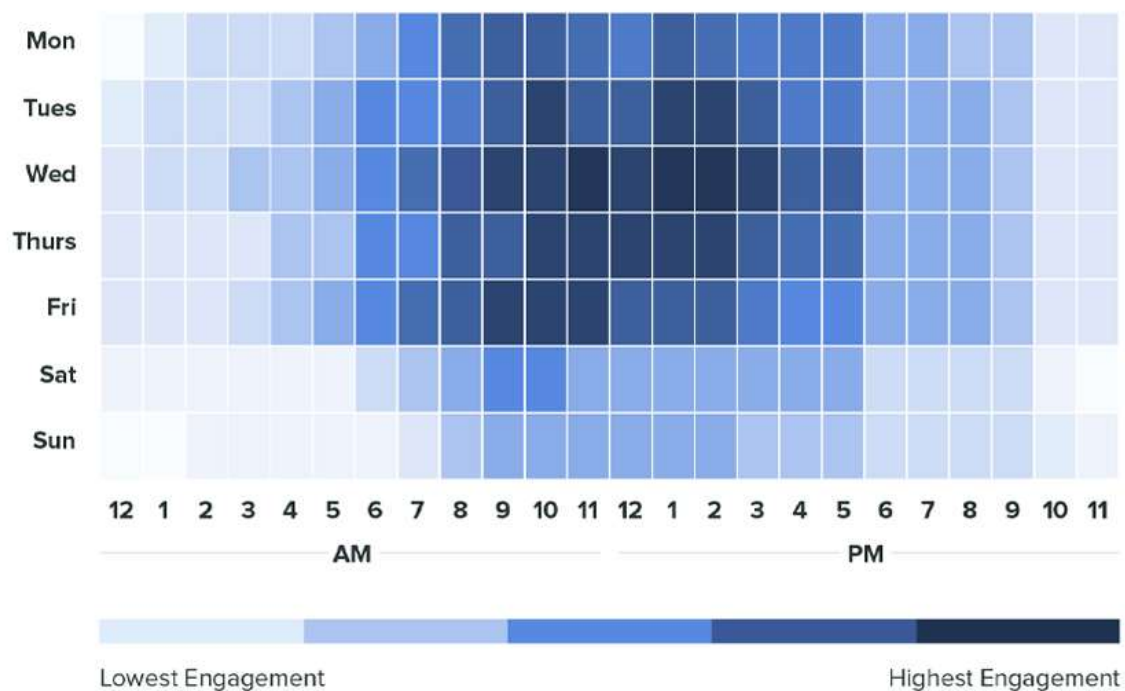
Assignment. No.2

Chart 5.2.1
Number of police officers in Crimeville, 2011 to 2019
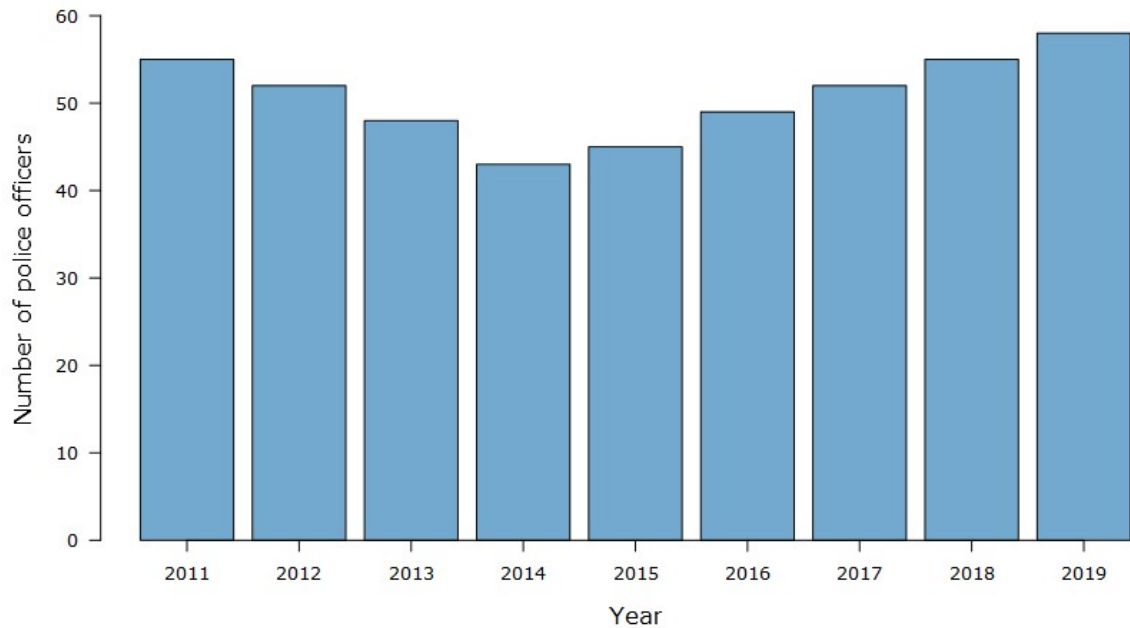
4. Heat Maps: Heat maps use color-coded cells in a matrix to represent data, where each cell's color intensity indicates the value of the data. They are useful for visualizing patterns, correlations, and distributions in large datasets, such as geographical data or gene expression data.

Assignment. No.2

Icon-Based Data Visualizations:

1. **Chernoff Faces:**
   Chernoff Faces is an iconic data visualization technique developed by Herman Chernoff in the 1970s. It uses facial features and expressions to represent multivariate data in a visual form. Each data point is represented as a face with multiple facial characteristics that correspond to the variables in the dataset.
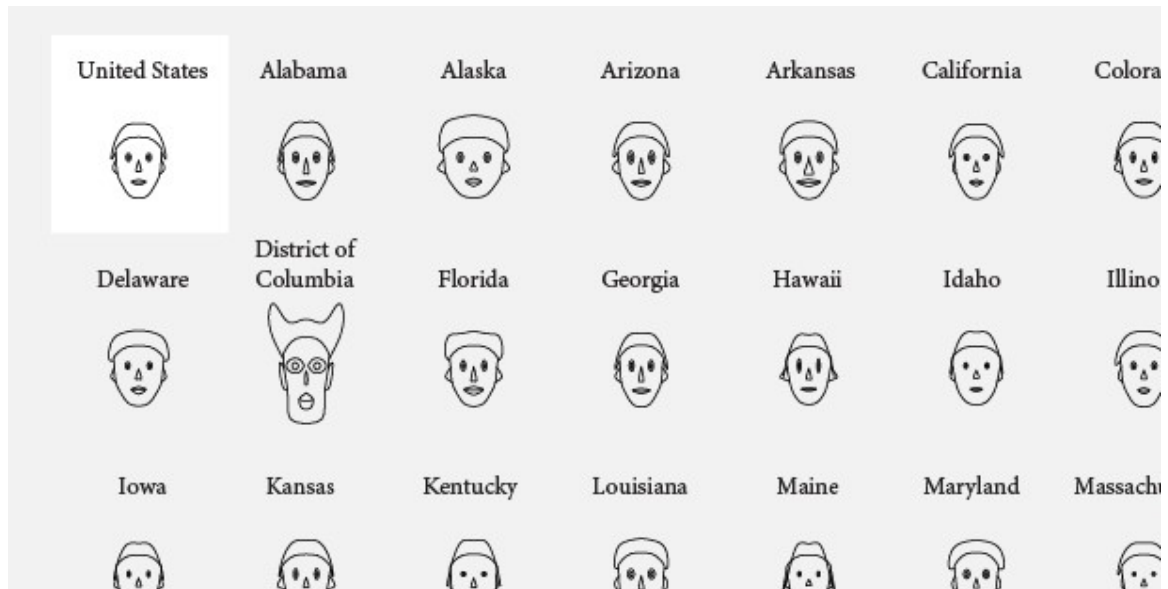
Key Features of Chernoff Faces:
- Variables are mapped to different facial features, such as the size of the eyes, shape of the mouth, length of the nose, etc.
- Each facial feature can represent a specific data attribute, such as age, income, education level, or any other quantitative or categorical variable.
- By observing the faces, analysts can visually identify patterns, correlations, and differences between data points.
- Chernoff Faces are especially useful for visualizing high-dimensional datasets where each data point has multiple attributes.

Advantages:
- Chernoff Faces offer a unique and memorable way to represent complex data, making them engaging and suitable for presentations and reports.
- They can effectively display relationships between variables and reveal hidden patterns that may not be apparent in traditional charts.

Disadvantages:
- Interpretation can be subjective and may vary among viewers.
- Drawing faces manually can be time-consuming and challenging for large datasets.
- May not be suitable for datasets with a large number of variables, as facial features can become cluttered and difficult to interpret.

2. **Gauge Chart:**
   Gauge charts, also known as dial or speedometer charts, are circular visualizations that resemble a traditional car speedometer or gauge. They are commonly used to represent a single value within a specified range or to show progress towards a goal or target.

Assignment. No.2

Sales Target & Sales Growth

Growth to last month
+3%

Difference to Target
+7%

Revenue vs Target - This Month
Revenue: $590,333

Revenue — Target

Key Features of Gauge Charts:
- The gauge is divided into intervals or sectors, each representing a specific range or category.
- A pointer or needle points to the value on the gauge, indicating the position of the data point within the range.
- Gauge charts are often used for measuring performance metrics, such as completion percentages, ratings, or levels of achievement.

Advantages:
- Gauge charts provide a simple and intuitive way to visualize a single value in comparison to a reference range or target value.
- They are visually appealing and can easily convey the status of a metric or progress towards a goal.

Disadvantages:
- Gauge charts are not suitable for displaying multiple data points or complex data relationships.
- They may not be the best choice for precise data comparisons, as the visual representation is limited to a single value within a range.