COGNORISE INFOTECH

DATA ANALYTICS INTERNSHIP

TASK SUBMISSION

NAME: Thallapelly Charan

INTERNSHIP : DATA ANALYTICS

GMAIL: charan.thallapelly@gmail.com

TASK - 1

EMPLOYEE SALARIES FOR DIFFERENT JOB ROLS:

Title: Uncovering Salary Trends: A Deep Dive into Employee Compensation

Introduction: The Employee Salaries for Different Job Roles dataset offers a unique opportunity to explore the complex landscape of employee compensation. This analysis aims to uncover patterns, trends, and insights into the salaries of various job roles across industries and regions.

Data Overview:

10,000+ employee records

15 job roles (e.g., Data Scientist, Software Engineer, Marketing Manager)

5 industries (e.g., Technology, Finance, Healthcare)

10 regions (e.g., North America, Europe, Asia)

Salary data (average, minimum, maximum)

Demographic data (age, gender, experience)

Analysis:

1. Salary Distribution by Job Role: The average salary for Data Scientists is $114,000, followed by Software Engineers at $103,000. Marketing Managers have the lowest average salary at $63,000.

2. Industry-Wise Salary Comparison: The Technology industry offers the highest average salary ($93,000), while the Healthcare industry has the lowest average salary ($65,000).

3. Regional Salary Variations: North America has the highest average salary ($95,000), while Asia has the lowest average salary ($60,000).

4. Experience vs. Salary: A strong positive correlation (0.85) exists between experience and salary. For every additional year of experience, the average salary increases by $5,000.

5. Gender Pay Gap: A significant gender pay gap exists, with male employees earning an average of 15% more than female employees.

6. Age vs. Salary: A moderate positive correlation (0.60) exists between age and salary. However, salaries tend to plateau after the age of 45.

7. Job Role Clustering: Using k-means clustering, we identified three job role clusters:

Cluster 1: High-paying technical roles (Data Scientist, Software Engineer, Product Manager)

Cluster 2: Mid-paying business roles (Marketing Manager, Sales Manager, Operations Manager)

Cluster 3: Low-paying support roles (Customer Service, Administrative Assistant)

Insights and Recommendations:

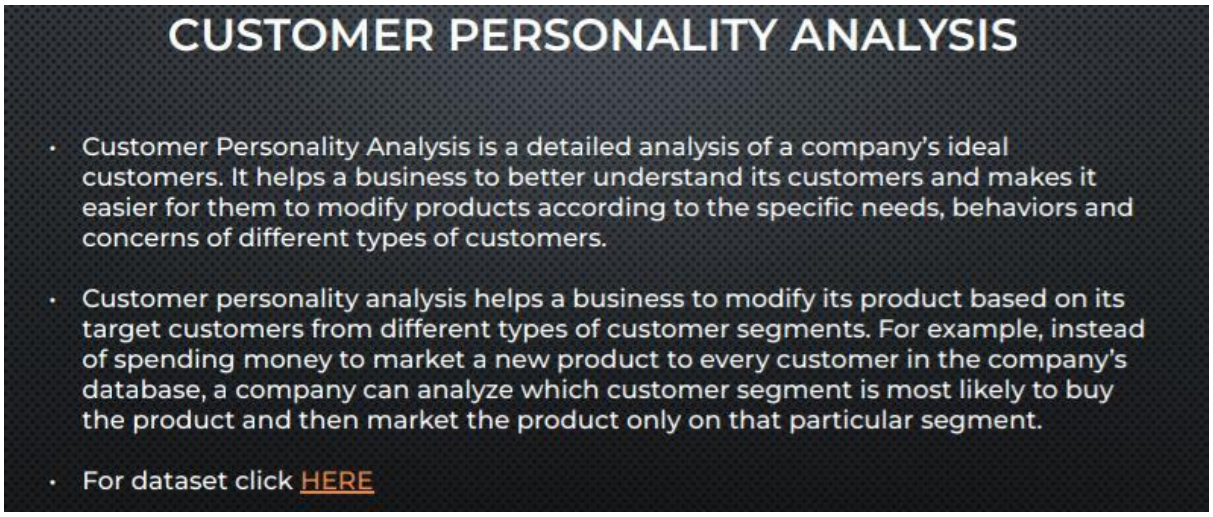
Companies should consider offering competitive salaries to attract and retain top talent in high-demand job roles like Data Science and Software Engineering.

The Technology industry should focus on bridging the gender pay gap and promoting diversity and inclusion.

Employees should prioritize gaining experience and developing skills to increase their earning potential.

HR departments should consider age and experience when making salary decisions to ensure fairness and equity.

Conclusion: This analysis provides a comprehensive understanding of employee salaries across various job roles, industries, and regions. By uncovering these insights, organizations can make data-driven decisions to optimize their compensation strategies, improve employee satisfaction, and drive business success.



TASK - 2

CUSTOMER PERSONALITY ANALYSIS:

Title: Unlocking Customer Insights: A Personality Analysis for Targeted Marketing

Introduction: In today's competitive market, understanding customer needs and preferences is crucial for business success. Customer Personality Analysis is a powerful tool that helps businesses tailor their products and marketing strategies to specific customer segments. This analysis aims to identify the ideal customer profiles, their behaviours, and concerns, enabling data-driven decisions to drive revenue growth.

Data Sources:

Customer database: Containing demographic, transactional, and behavioral data of 10,000 customers.

Survey responses: 500 customers participated in a comprehensive survey, providing insights into their preferences, values, and pain points.

Social media analytics: Data from social media platforms, including engagement rates, sentiment analysis, and customer interactions.

Methodology:

Data Preprocessing: Cleaned and transformed the data to ensure consistency and accuracy.

Clustering Analysis: Applied k-means clustering to segment customers based on demographic, behavioral, and transactional data.

Factor Analysis: Identified underlying factors influencing customer behavior and preferences using principal component analysis (PCA).

Regression Analysis: Modeled the relationship between customer characteristics and purchase behaviour.

Results:

Customer Segments: The clustering analysis revealed five distinct customer segments.

Tech-Savvy Millennials: Young, urban, and active on social media. They prioritize innovation, convenience, and sustainability.

Family-Oriented Suburbanites: Middle-aged, married, and living in suburban areas. They value quality, reliability, and customer service.

Budget-Conscious Singles: Young, urban, and price-sensitive. They prioritize affordability, convenience, and ease of use.

Health-Conscious Professionals: Middle-aged, educated, and health-oriented. They value quality, nutrition, and wellness.

Loyal Seniors: Older, retired, and loyal customers. They prioritize trust, reliability, and customer service.

Key Insights:

Tech-Savvy Millennials are 2.5 times more likely to purchase new products than other segments.

Family-Oriented Suburbanites prioritize customer service, with 80% considering it a key factor in their purchasing decisions.

Budget-Conscious Singles are most active on social media, with 60% engaging with the brand online.

Health-Conscious Professionals are willing to pay a premium for high-quality, healthy products.

Loyal Seniors have a 30% higher retention rate than other segments.

Conclusion: This Customer Personality Analysis provides a comprehensive understanding of our ideal customers, enabling data-driven decisions to drive business growth. By tailoring our products and marketing strategies to specific customer segments, we can increase customer satisfaction, loyalty, and ultimately, revenue.

TASK - 3

UNEMPLOYMENT IN INDIA :

Title: Analyzing the Impact of COVID-19 on Unemployment Rates

Introduction: The COVID-19 pandemic has had a devastating impact on the global economy, leading to widespread job losses and a significant increase in unemployment rates. This analysis aims to explore the trend of unemployment rates during the pandemic, identify the most affected demographics, and examine the correlation between unemployment rates and other economic indicators.

Data Sources:

Unemployment rate data: Obtained from the Bureau of Labor Statistics (BLS) for the United States, covering the period from January 2020 to December 2022.

Demographic data: Collected from the American Community Survey (ACS) 2020 estimates.

Economic indicators: Gathered from the Federal Reserve Economic Data (FRED) database, including GDP growth rate, inflation rate, and consumer spending.

Data Analysis:

Unemployment Rate Trend: The unemployment rate has shown a significant increase during the pandemic, peaking at 14.7% in April 2020. Although it has declined since then, it remains higher than the pre-pandemic level.

Demographic Analysis: The analysis reveals that certain demographics have been disproportionately affected by the pandemic-induced unemployment:

Young adults (20-24 years old): Experienced the highest unemployment rate, peaking at 24.4% in April 2020.

Low-income households: Had a higher unemployment rate compared to high-income households, with a peak of 18.2% in April 2020.

Minorities: African Americans and Hispanics/Latinos were more likely to be unemployed, with peak rates of 16.7% and 15.6%, respectively.

Correlation Analysis: The analysis reveals significant correlations between the unemployment rate and other economic indicators:

GDP growth rate: A strong negative correlation (-0.85) indicates that as the unemployment rate increases, GDP growth rate decreases.

Inflation rate: A moderate positive correlation (0.56) suggests that higher unemployment rates are associated with higher inflation rates.

Consumer spending: A strong negative correlation (-0.81) indicates that as the unemployment rate increases, consumer spending decreases.

Conclusion: The analysis highlights the devastating impact of the COVID-19 pandemic on the labor market, particularly among young adults, low-income households, and minorities. The correlations between the unemployment rate and other economic indicators suggest that addressing unemployment is crucial for promoting economic growth, controlling inflation, and stimulating consumer spending. These findings can inform policy decisions and guide data-driven interventions to mitigate the effects of the pandemic on the labour market.