

Tools and Techniques in Data Science

Assignment

Problem#01:

Perform an in-depth EDA on the Titanic dataset. Focus on the following tasks:

- Handle missing data appropriately, considering imputation strategies.
- Create new features or extract useful information from existing features.
- Conduct advanced grouping and aggregation to analyze survival rates based on multiple factors simultaneously.
- Visualize the results using Matplotlib or Seaborn for effective communication.

Problem#02:

Use the IMDb movies dataset for exploratory data analysis. Perform the following tasks:

- Analyze the distribution of movie genres over time.
- Identify the top-rated directors and actors.
- Explore the correlation between budget, revenue, and ratings.
- Utilize Pandas for advanced indexing and filtering.

Problem#03:

Conduct an EDA on the World Happiness Report dataset. Focus on the following aspects:

- Analyze the distribution of happiness scores and identify countries with significant changes over time.
- Investigate correlations between different factors influencing happiness.
- Perform a regional analysis to identify patterns or outliers.
- Utilize advanced grouping and aggregation techniques for insightful summaries.