

Assignment Objective

Analyze accident data to extract actionable insights on accident severity and test hypotheses related to temporal, weather, and visibility factors affecting accidents.

Assignment Tasks

1. Formulate Analytical Questions

Develop 3 to 5 hypothesis-driven questions related to accident severity. Example questions could include:

- Do accidents during nighttime have higher severity than daytime?
- Is low visibility associated with increased accident severity?
- Does the road surface condition impact accident severity distribution?

2. Data Filtering and Grouping

Use pandas to filter and group data by the features relevant to your hypotheses. Use aggregation functions to summarize severity counts or averages.

3. Statistical Hypothesis Testing

For each question, perform an appropriate statistical test (t-test, chi-square test, ANOVA, etc.) to validate or reject the hypothesis. Clearly state:

- Null and alternative hypotheses
- Significance level (α)
- Test statistics and p-values
- Conclusion based on p-value

4. Visualize Results

Create visualizations that support your analysis. Use bar charts, box plots, or heatmaps to depict patterns behind the significance tests.

Ensure charts have titles, labels, and legends for clarity.

5. Document Assumptions and Limitations

Clearly document assumptions made during the analysis, such as independence of observations or normality assumptions. Discuss potential limitations or biases in the dataset.

6. Summarize Key Insights

Write a concise summary of findings focusing on how the factors analyzed affect accident severity and implications for road safety.

Deliverables

- Python script or Jupyter Notebook with:
 - Data filtering and grouping code
 - Statistical tests implementation and results
 - Visualizations aligned with hypotheses
 - Markdown cells or comments explaining methodology and insights
- A short report (1-2 pages) summarizing hypotheses, tests, results, and recommendations.