Assignment 4 Becoming an Independent Data Scientist Example Solution

1 Region and Domain

State the region and the domain category that your data sets are about.

Cape Town, UK

Passenger information

2 Research Question

You must state a question about the domain category and region that you identified as being interesting.

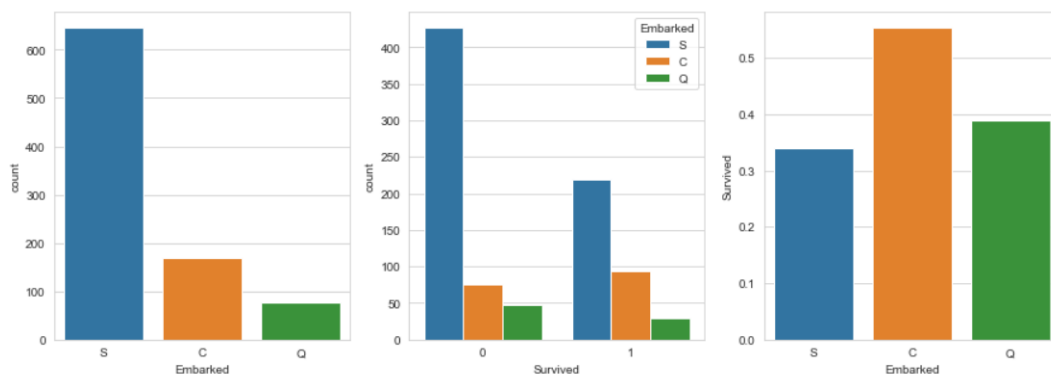Whether the embarked places had influence on the survivorship?

3 Links

You must provide at least two links to publicly accessible datasets. These could be the links to files such as CSV or Excel files, or links to websites which might have data in tabular form, such as Wikipedia pages.

https://www.kaggle.com/c/titanic/data

https://www.kaggle.com/c/titanic/overview

4 Image

You must upload an image which addresses the research question you stated. In addition to addressing the question, this visual should follow Cairo's principles of truthfulness, functionality, beauty and insightfulness

5 Discussion

You must contribute a short (1-2 paragraph) written justification of how your visualization addresses your stated research question.

This visualization was concerned with answering the question of how the embarked location influenced on the survivorship. The counts of embarked locations, the counts of survivors based on embarked location and the survivorship based on embarked location were plotted. The plot indicates that the passengers embarked from C-Cherbourg had an evident higher survivorship than other locations (Q = Queenstown, S = Southampton).