

A Poster Presentation of IPE-205 on “Haberman’s survival : Exploratory Data Analy f7f288”

DATASET

A	B	C	D
30	64	1	1
30	62	3	1
30	65	0	1
31	59	2	1
31	65	4	1
33	58	10	1
33	60	0	1
34	59	0	2
34	66	9	2
34	58	30	1

Code

```
# check for the input dataset
import os
print(os.listdir('../input'))
```

```
# import necessary packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

```
# load the dataset
cancer_df =
pd.read_csv('../input/haberman.csv',
header=None, names=['age',
'year_of_treatment',
'positive_lymph_nodes',
'survival_status_after_5_years'])
print(cancer_df.head())
```

```
print(cancer_df.info())
```

```
fig, axes = plt.subplots(1, 3,
figsize=(15, 5))
for idx, feature in
enumerate(list(cancer_df.columns):
-1)):
sns.violinplot(
x='survival_status_after_5_years',
y=feature, data=cancer_df,
ax=axes[idx])
plt.show()
```

Pair plot in seaborn plots the scatter plot between every two data columns in a given dataframe.

It is used to visualize the relationship between two variables

```
sns.pairplot(cancer_df,
hue='survival_status_after_5_years', size=4)
plt.show()
```

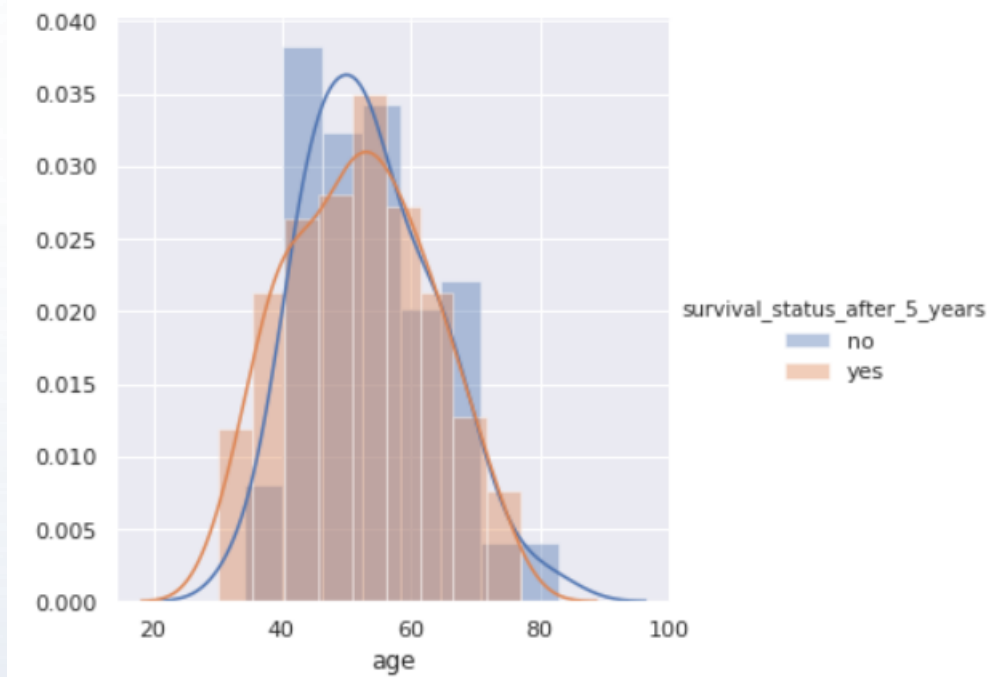
```
# modify the target column values to be
meaningful as well as categorical
cancer_df['survival_status_after_5_years'] =
cancer_df['survival_status_after_5_years'].map({1:
"yes", 2:"no"})
cancer_df['survival_status_after_5_years'] =
cancer_df['survival_status_after_5_years'].astype('
category')
print(cancer_df.head())
```

```
for idx, feature in
enumerate(list(cancer_df.columns):
-1)):
fg =
sns.FacetGrid(cancer_df,
hue='survival_status_after_5_
years', size=5)
fg.map(sns.distplot,
feature).add_legend()
plt.show()
```

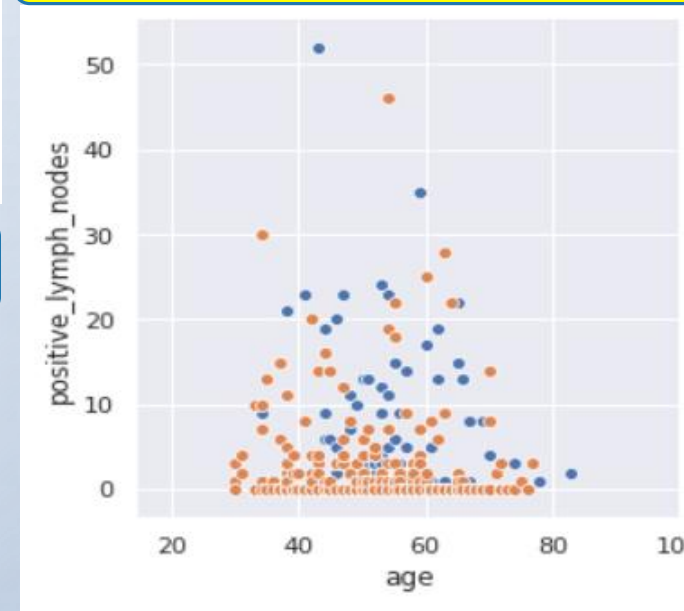
```
# print the unique values of the
target column
print(list(cancer_df['survival_st
atus_after_5_years'].unique()))
```

Summary of the Poster:
The poster is on Haberman’s survival-Exploratory Data Analysis. It represents some statistical graphical figures on this dataset. It has 306 rows and 4 columns.

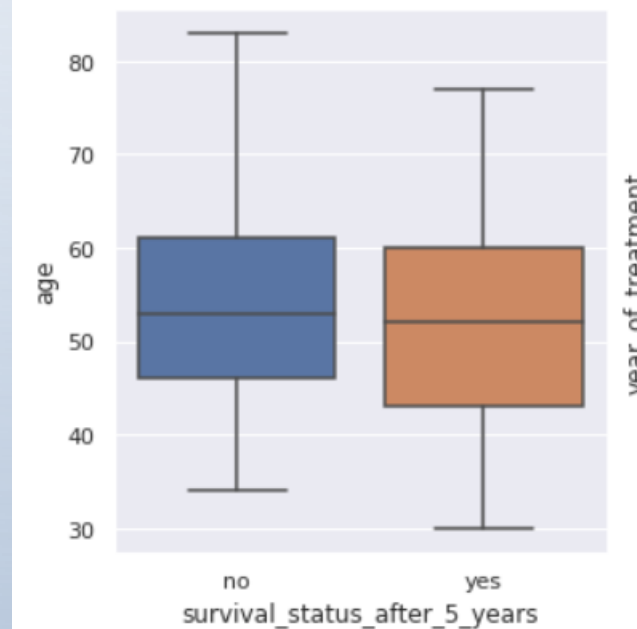
Distribution Plot



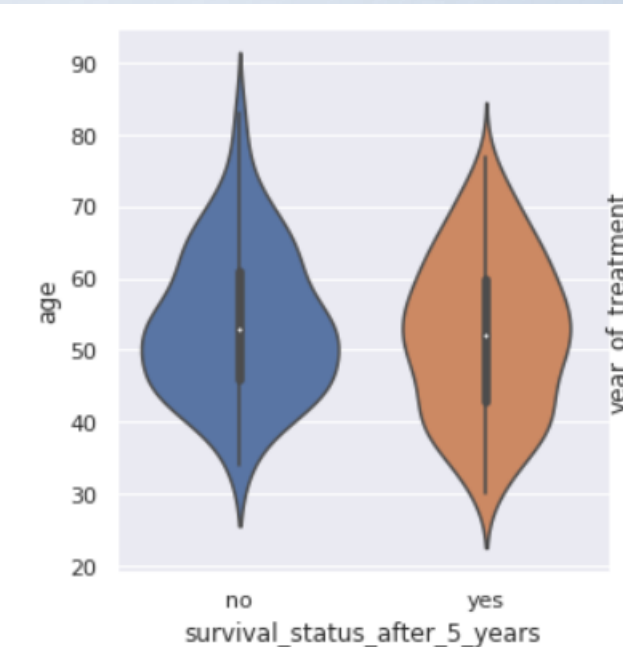
Scatter Plot



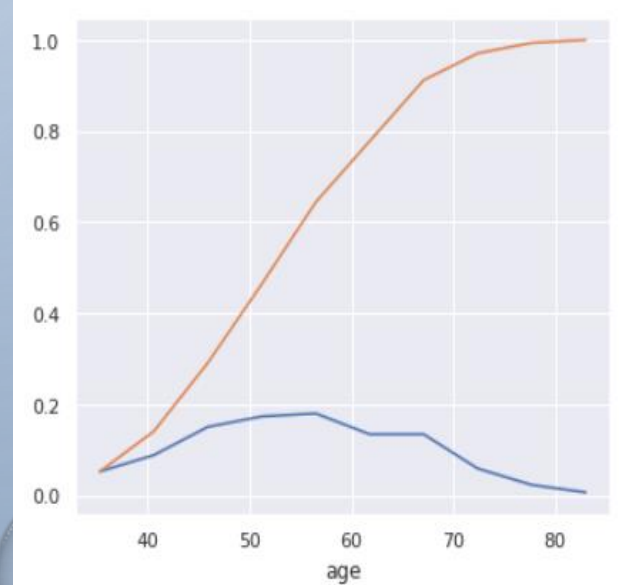
Box Plot



Violin Plot



Graphical Representation



Submitted to:

Tanmoy Das

Submitted by:

MD. Tanzil Islam
Roll: 201736021

Fahim Montasir
Roll: 201736022

Syed Musfiqur Rahman
Roll: 201736047

Nahian Ibtada
Roll: 201736027

Kaggle link: <https://www.kaggle.com/tanzilislam021/haberman-s-survival-exploratory-data-analy-f7f288/edit>
Github link: <https://github.com/mdtanzilislam/Poster>