Exploratory Data Analysis Retail ~ GRIP

d1= given data set in .csv file

d2= data set sorted by the 'negative' profit column

b1= top 20 values, sorted by sales

A) Steps~

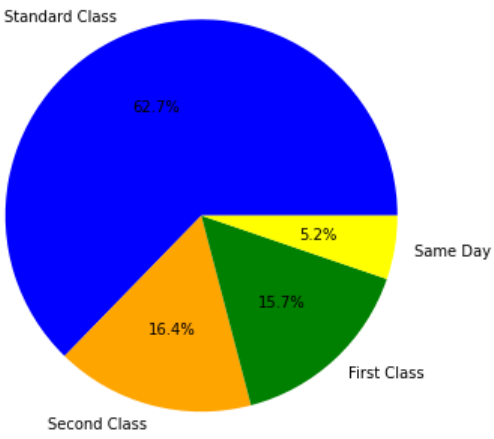
- CSV read through ***pandas*** library
- Analysing the dataset from scratch

- **9994 total records**
- **1871 records with negative profit**
- **4 different shipping modes namely, Standard Class, First Class, Second Class, Same Day**
- **13 columns**

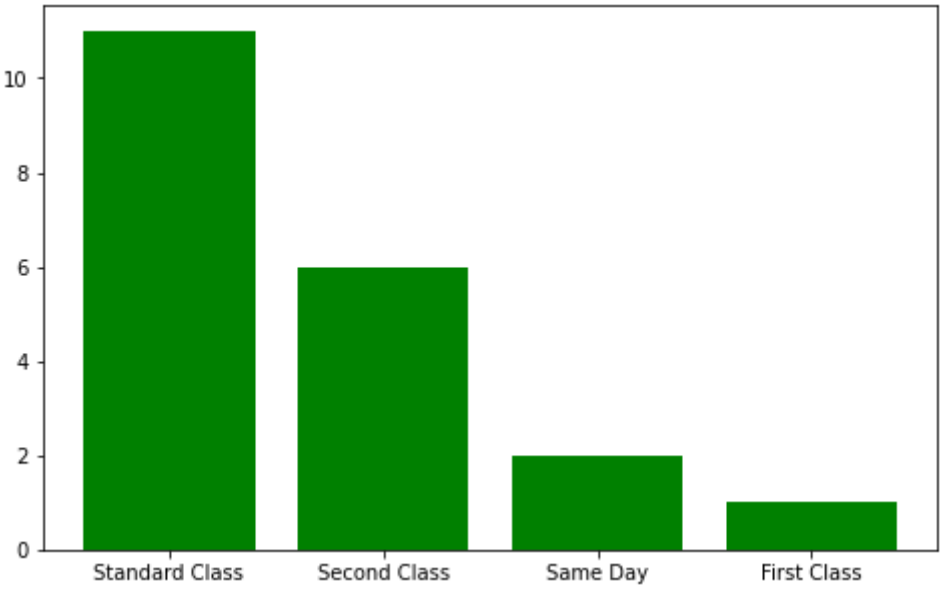
B) Weak Areas & Business Problems~

- Profit lies between **-6599 to -0.09**
- Percentages for different Ship Modes:

a) **Standard Class~** 62.7%
b) **Second Class~** 16.4%
c) **First Class~** 15.7%
d) **Same Day~** 5.2%

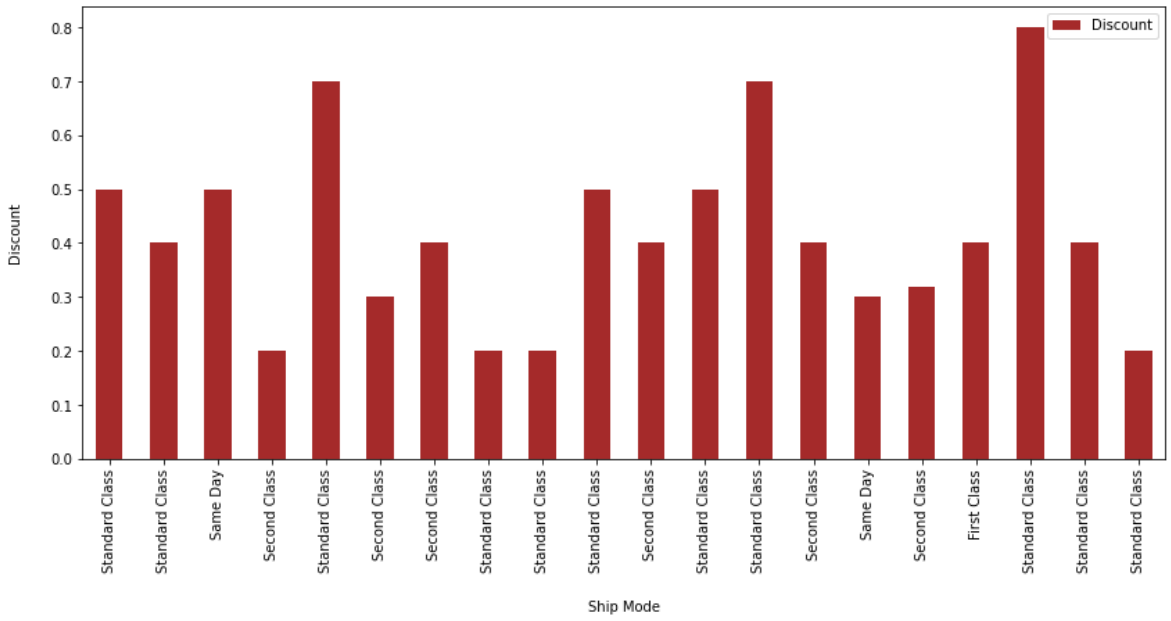


Percentage of Ship Modes used



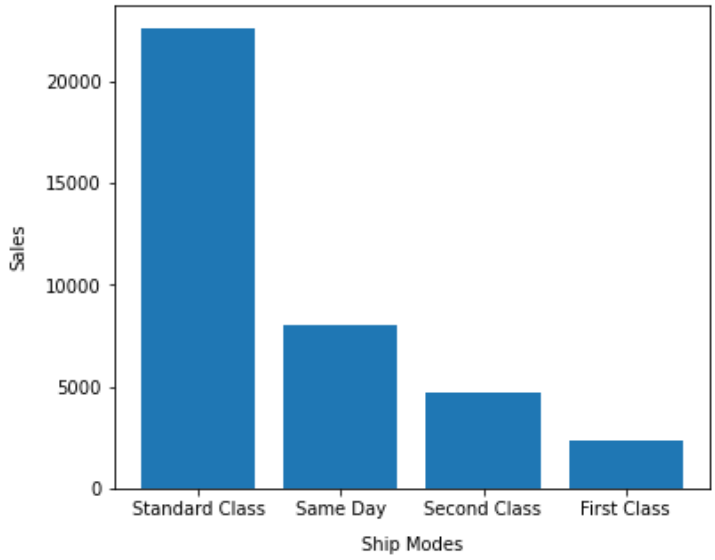
Ship Modes & their count (Top 20 values)

- Heavy discount between 20% to 80% when the Ship Mode is **"Standard Class"** for top 20 values.
- Maximum loss is of **-6599** for **Standard Class** while looking at overall data with **70% discount**.



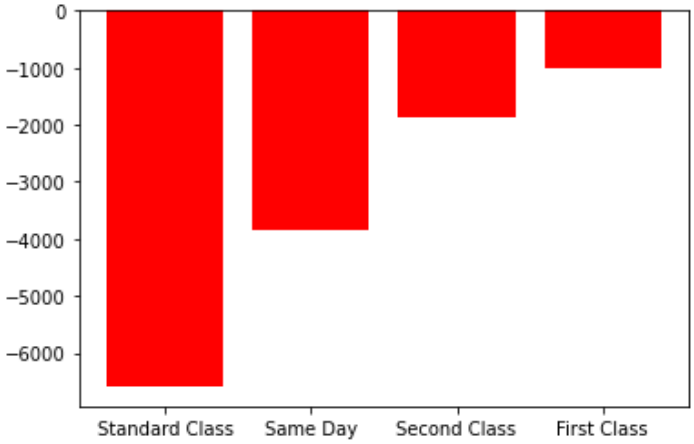
Ship Modes and Discount

- Minimum loss is of **-25.05** for **Standard Class** while analysing top 20 records with **20% discount**.
- Maximum profit is of **8399.98** for Standard Class while looking at complete dataset with **0% discount**.
- Maximum sales is of **22638.48** when the Ship Mode is **Standard Class**, in overall dataset.



Ship Mode & Sales

- Minimum sales is of **2003.92** for Standard Class in top 20 values.
- Minimum sales is of **0.44** for Standard Class in overall records.
- Maximum Loss is of -6599 in Standard Class



Maximum Loss

Solutions:

- Discount needs to be reduced below 20% for **“Standard Class”** so that profit can be maximized.