

Day1

1. Assignment

Use the following link and answer the following questions.

<https://raw.githubusercontent.com/justmarkham/pandas-videos/master/data/chipotle.tsv>

1. What is total item price.
 2. What is minimum, maximum, average, 25% and 75% of item price
 3. Display all rows with item_name starts with "Chips"
2. Use tips dataset from seaborn, and display graph to show male and female smokers and non smokers
- a. Find daywise min, max, average, and 25 and 75 percentile of total bill
 - b. On every data how many males and how many females go to hotel
 - c. On which day number of visitors are maximum and on which day the number of visitors are min.

Day 3

1. Using following equation generate data for polynomial regression.

$$X = 6 * \text{np.random.rand}(200, 1) - 3$$
$$y = 0.7 * X^{**3} + 0.8 * X^{**2} + 0.85 * X + 2 + \text{np.random.randn}(200, 1)$$

2. Find accuracy score for degree 2, 3, 4, 5, 6, 7, 8 for above data, display best suitable degree for the scenario, also display coefficients.
3. Use **hiring.csv**. This file contains hiring statics for a firm such as experience of candidate, his written test score and personal interview score. Based on these 3 factors, HR will decide the salary. Given this data, you need to build a machine learning model for HR department that can help them decide salaries for future candidates. Using this predict salaries for following candidates,

2 yr experience, 9 test score, 6 interview score

12 yr experience, 10 test score, 10 interview score

Answer

53713.86 and 93747.79

4. Predict canada's per capita income in year 2020. There is an exercise folder here on github at same level as this notebook, download that and you will find `canada_per_capita_income.csv` file. Using this build a regression model and predict the per capita income for canadian citizens in year 2020

Answer

41288.69409442