**DSC Study pods 2.0**

**Study pods no. –** SP0232

**Technology :-** Data Science

**Team Members-** Tanya Agrawal, Lokesh Bhati, Anurag Gupta, Tanmay Goyal

**Mentor**:- Mr. Shadab Hussain

**DESCRIPTION:-**

We have taken data set from Kaggle **To predict whether customer will purchase given travel package or not.**

"Trips & Travel.Com" company wants to enable and establish a viable business model to expand the customer base. There are 5 types of packages the company is offering - ***Basic, Standard, Deluxe, Super Deluxe, King.*** Therefore, We pitch a package to customer and predict whether they would purchase it or not.

Looking at the data of the last year, we observed that 18% of the customers purchased the packages. However, the marketing cost was quite high because customers were contacted at random without looking at the available information.

For this ,customer care of company contacted to customer and pitch a travel package & get feedbacks whether they like the products or they would purchase the products according to their background condition i.e age , monthly income, marital status ,desination etc or according to their requirements.

**For Prediction,**

We have taken a random data i.e Test Data and make prediction over this by making a record of all the statistics by contacting the customers and handed over this data to company.

**Description of Files and Data:-**

We have total of 6 files :- 3 .csv files and 2 .ipynb files and 1 .joblib file

In the main file, we have 4888 rows and 16 columns after data visualisation i.e after removing all the duplicates and null values.

In train\_data\_prediction.csv, there are 3666 rows and 16 columns

In test\_data\_prediction.csv, there are 1222 rows and 16 columns

**Resuources:-**

Statistics:

<https://www.khanacademy.org/math/statistics-probability/summarizing-quantitative-data>

Data Distribution:

<https://www.khanacademy.org/math/statistics-probability/modeling-distributions-of-data>

Probability:

<https://www.khanacademy.org/math/statistics-probability/probability-library>

[https://ocw.mit.edu/courses/mathematics/18-05-introduction-to-probability-and-statistics-spring-2014/](https://www.khanacademy.org/math/statistics-probability/probability-library)

Linear Algebra:

<https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab>

[https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/](https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab)

Linear Regression:

<https://towardsdatascience.com/introduction-to-linear-regression-and-polynomial-regression-f8adc96f31cb>

<https://medium.com/swlh/understanding-mathematics-behind-normal-equation-in-linear-regression-aa20dc5a0961>

<https://towardsdatascience.com/introduction-to-linear-regression-and-polynomial-regression-f8adc96f31cb>

Machine Learning Lectures by Andrew Ng:

<https://www.youtube.com/playlist?list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU>

[https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-867-machine-learning-fall-2006/](https://www.youtube.com/playlist?list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU)

Calculus:

<https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr>

[https://www.khanacademy.org/math/calculus-1](https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr)

**Experience-**

It was really nice experience participating in study pods .We got to explore a new field and learnt a lot of new concepts. We learnt python, Probability, Statistics, Basic Algorithms of ML, Regression (Linear Regression, logistic Regression) and many more. All team members and mentor were very supportive.