



Home Work

Data Science & Machine Learning

Data Analysis

NUMBER	DEADLINE	TOPICS
Optional_1	Thursday, 1402/11/26 23:59:59 PM	Data Analysis, Pandas, Data Visualization

۱. پاسخ تمرین را در کارپوشه خود، ارسال و پشتیبان فنی @Daneshkar_mentor را منشن کنید

۲. جواب سوالها را با فرمت ipynb در یک فایل فشرده تجمیع و ارسال کنید.

۳. در صورتی که سوال و یا ابهامی دارید در گروه چت تلگرامی بپرسید.

۴. لطفا نام فایل ارسالی خود را به صورت YourName_HWOp1 با نام خود قرار دهید

Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate? What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests?

Use the attached dataset and explore and analyze the data to discover important factors that governs booking.

Data Cleaning:

As a first step, **explore the data and preprocess it in order to be cleaned:**

1. Check for missing values and find the best practice to deal with them
2. Convert the dtypes to appropriate ones according to the values
3. Check for the duplicate values and remove them from data
4. According to the columns and their meanings, do any other preprocessings necessary for further analysis

Analysis:

1. There are two types of hotels in the dataset, which type of hotels booked most? Plot a pie chart to show the results.
2. What percentage of bookings were cancelled? What type of chart do you suggest to show the results?
3. Based on different hotel types, what percentage of bookings were cancelled? What type of chart do you suggest to show the results?
4. In different years, what type of hotels were booked **most** and booked **least**? Draw a proper countplot to show the results.
5. According to different hotel types, how 'adr' changes in different

months?

6. Check hotel wise daily rates mean, standard deviation, median, min and max.
7. Extract the number of bookings per country and investigate that most of the guests are from which countries?
8. According to the 'market_segment' most of the guests book their rooms on what channels?
9. What other useful analysis can you extract?

Notes:

Write proper markdowns according to your results and include the highlighting points from each analysis

Happy Analyzing 😊