PROJECT-2
DATA HARMONIZATION AND INSIGHTS EXTRACTION

PRESENTED BY:-ASHISH KUMAR SEN (NEXT HIKE IT SOLUTION)

OBJECTIVES

PERFORM EXPLORATORY DATA ANALYSIS AND VISUALIZE THE DATA TO UNDERSTAND THE ENVIRONMENTAL AND SEASONAL SETTINGS

PREDICT BIKE RENTAL COUNTS BASED ON ENVIRONMENTAL AND SEASONAL SETTINGS WITH THE HELP OF A PYTHON LIBRARIES

PREREQUISITES

EXPLORATORY DATA ANALYSIS

DATA MANIPULATION

DATA VISUALIZATION

PYTHON

MODULES AND LIBRARIES

INDUSTRY RELEVANCE

EXPLORATORY DATA ANALYSIS: IT FINDS TRENDS AND PATTERNS, OR CHECKS ASSUMPTIONS BY ANALYZING DATA WITH VISUAL TOOLS

DATA MANIPULATION : IT ORGANIZES AND CHANGES INFORMATION TO MAKE IT MORE UNDERSTANDABLE.

DATA VISUALIZATION : ITREPRESENTS DATA WITH THE USE OF COMMON GRAPHS, PLOTS, OR CHARTS.

PANDAS ,MATPLOTLIB, SEABORN, NUMPY,: IT IS USED FOR STATISTICAL ANALYSIS, GRAPHICS REPRESENTATION, AND REPORTING.

PYTHON: IT HELPS SOFTWARE PROGRAMS IN BEING MORE ACCURATE AT PREDICTING OUTCOMES WITHOUT EXPLICITLY PROGRAMMING THEM TO DO SO.

DATASET DESCRIPTION

VARIABLE-DESCRIPTION

INSTANT-RECORD INDEX

DTEDAY - DATE

SEASON -(SEASON - 1 = SPRING, 2 = SUMMER, 3 = FALL, 4 = WINTER)

YR- YEAR (2011)

MNTH- MONTH (1 TO 12)

HOLIDAY-WEATHER DAY IS A HOLIDAY OR NOT

WEEKDAY- DAY OF THE WEEK

WORKING DAY - WHETHER THE DAY IS NEITHER A WEEKEND NOR A HOLIDAY ETC.

TASKS TO PERFORM

PERFORM THE FOLLOWING TASKS ON THE DATASET PROVIDED USING PYTHON

EXPLORATORY DATA ANALYSIS : LOAD THE DATASET AND THE RELEVANT LIBRARIES

PERFORM DATA TYPE CONVERSION OF THE ATTRIBUTES

CARRY OUT THE MISSING VALUE ANALYSIS

ATTRIBUTES DISTRIBUTION AND PLOT MONTHLY DISTRIBUTION OF THE TOTAL NUMBER OF BIKES RENTED

PLOT YEARLY DISTRIBUTION OF THE TOTAL NUMBER OF BIKES RENTED

PLOT BOXPLOT FOR OUTLIERS' ANALYSIS

TASKS TO PERFORM

SPLIT THE DATASET INTO TRAIN AND TEST DATASET

PREDICT THE PERFORMANCE OF THE MODEL ON THE TEST DATASET

PROJECT OUTCOME

THIS PROJECT IS DESIGNED TO HELP UNDERSTAND HOW TO PERFORM EXPLORATORY DATA ANALYSIS, PLOT GRAPHS, AND PREDICT USING A PYTHON ALGORITHM.

- YOU SHOULD BE ABLE TO ANALYZE THE DATASET FOR THIS PROJECT TO CREATE A REPORT. YOU WILL BE ABLE TO USE A PYTHON ALGORITHM AND PREDICT THE BIKES RENTED DAILY.

COMPLETE EACH TASK LISTED IN THE PROBLEM STATEMENT

THANK YOU