

MAR657: Visual Analytics

Individual Assignments 6 & 7: Presenting Results from Hypothesis Testing and Data Analysis

DATA SET: 'SIVA.CSV'

SUBMISSION: A **PDF** FILE

The submission should be formatted in the following way: FirstNameLastName_Individual Visual Analytics Assignment ##.pdf or FirstNameLastName_Individual Visual Analytics Assignment ##.R) (**Example.** EmilyKo_Individual Visual Analytics Assignment 1.pdf)

Due date: 11:59 PM ET on October 21st (Tuesday)

- 1) Create **ONE correlation plot**, **ONE visualization for a regression model**, and **ONE visualization for a cluster model**. Please refer to the R script that was discussed at the class meeting. In order to create visualizations for statistical models, you have to run the model first (e.g., regression model, clustering model).
 - For the regression model, "Recom_mend_Siva" is highly recommended as a dependent variable.
 - For the correlation analysis and the cluster analysis, all variables that you are running **SHOULD** be numerical variables. Followings are the examples of the numerical variables in the data set: 'Staff_Courtesy', 'Speed_of_Service', 'Veh_Equip_Condition,' but there are other numerical variables other than these so please check the data types using `str` function in R.
 - You may need to subset or remove NA values in order to run the models or create visualizations. In that sense, the following syntax can be considered:

Subset :

```
subset <- siva[,c(C1,C2,C3,C4)]
```

Remove NAs :

```
subsetScale2 <- na.omit(siva)
```

SIVA Data

This data set includes transactions of Siva rental cars in those locations in the U.S. and Canada and customer satisfaction surveys for those transactions. Customer satisfaction survey responses linked to car rental transactions, for randomly chosen customers.

Variables

Column Name	Data Type	Human Notes
xgra_n1clb_nbr	String	Customer ID: Club number of customer if customer is part of Club
Siva_Rental_Number	String	Unique identifier of each rental transaction
Rent_area_loc	String	Code for rental location
Date_of_Survey	Datetime	The date on which the survey was taken
Day_of_Week	String	Day of the week when survey was taken
Time	Time	Time of day that the rental transaction started
Survey_Type	String	Method used by customer to complete survey
Purpose_of_Rental	String	All values are: Leisure/Personal, Business or Insurance Replacement/Loaners
Recom_mend_Siva	Integer	Score of 0=Not at All Likely to 9=Ext. Likely; question as it appears on survey instrument reads: "How likely is it that you would recommend Siva to a friend or colleague?"
Staff_Courtesy	Integer	How courteous the staff was, on scale of 0=not courteous at all to 9=extremely courteous. Survey instrument reads: "Please rate your experience with us in the following areas: Courtesy of staff"
Speed_of_Service	Integer	Speed of service, on a scale of 0=Extremely slow to 9=Extremely fast, Survey instrument reads: "Please rate your experience with us in the following areas: Speed of Service"
Veh_Equip_Condition	Integer	On a scale of 0=Poor, 9=excellent, the condition that the vehicle was in at the time of rental, Survey instrument reads: "Please rate your experience with us in the following areas: Condition of vehicle & equipment"
Trans_Billing_as_Expected	Integer	On a scale of 0=Poor to 9=Excellent, survey instrument reads: "Please rate your experience with us in the following areas: Transaction and/or Billing as Expected"
Value_for_the_Money	Integer	On a scale of 0=Poor to 9=Excellent, Survey Instrument reads: "Please rate your experience with us in the following areas: Value for the money"
Area	String	A code and description of the location
loc_nm	String	Location name for rental location
ga_region_desc	String	Siva administrative division of the location where the car was rented. NOTE: CA LICENSEES (less than 1% of transactions) are California businesses with the right to use the Siva name

xgra_ckot_ts	Datetime	Date of vehicle checkout
xgra_ckin_ts	Datetime	Date of vehicle return
xgra_vclass_reserv	String	Vehicle class that customer reserved
xgra_veh_class	String	Vehicle class that customer received/drove
xgra_birth_date	Date	Birth date of customer
rent_loc_type	String	Location Type: 51% of surveys are associated with AP (airport), 47% of surveys are associated with OFF AP (off airport), small percentage: other. Breakdown of Locations: 79% of all locations are OFF AP, 20% of all locations are AP, small percentage: other.
cust_tier_code	String	Alphanumeric code that describes the loyalty program membership tier of customer, vast majority are "NA" (no level) or "RG" (basic).
booking_channel_code	String	Retail channel through which the reservation was made.
col34_total_charges	Numeric	Total actual bill paid
col38_currency	String	Indicates if the customer paid in USD or CAD
Total_Charge USD	Numeric	Total Rental Charge in USD
Survey_checkout_diff	Numeric	Number of days between checkout and taking the survey
booking Channel Dummy	Numeric	=1 if booking was done through Siva.com =0 otherwise