T 201: Experimentation and Uplift Testing This notebook serves as the solution of the second task Experimentation and Uplift Testing of Quantium Data Analytics Virtual **Experience Program.** In the previous task we conducted analysis on the client's transaction dataset and identified customer purchasing behaviours to generate insights and provided commercial recommendations. We now extend our analysis from Task 1 to help us identify benchmark stores that allow you to test the impact of the trial store layouts on customer sales. Category Manager for Chips, has asked us to test the impact of the new trial layouts with a data driven recommendation to whether or not the trial layout should be rolled out to all their stores. They new trial layouts were rolled in three stores, store number 77, 86, and 88 during trial period. The trial period was for three months, i.e. February, March, and April 2019. We will complete this task in two phases. • Phase 1: We will find control store corrosponding to the trial stores • Phase 2: Assessment of trial stores during the trial period What is Control Store? Store that have similar performance to trial store in terms of measure like monthly sales, number of customers, etc. Sequence of Analysis 1. We will be considering *monthly sales and number of customers* to find the potential control store. 1. Then we will define driving metrics to rank the stores on how similar they are to the trial store. · Correlation Score · Absolute Distance Score 1. We will select the store having higest score as control store for the corrosponding trial store. • Only pre-trial period will considered to select the control store We will also visualise the performance to cross check that we have selected right control store 1. Once we get control store for each trial store, we will assess the trail store performance during the trial period. We will use statistical tests and visualizations to test the impact of the new trial layouts for these stores 1. Finally, we will make conclusion on whether or not the trial layout should be rolled out to all their stores. **Importing Libraries** In [1]: import numpy as np import pandas as pd import matplotlib.pyplot as plt import seaborn as sns from scipy import stats sns.set(font scale=1) import warnings warnings.filterwarnings("ignore") **Loading Dataset** In [2]: | #loading dataset df = pd.read\_csv("C:/Users/amitm/Jupyter Notebools/Quantium/Task 2/QVI\_data.csv") #checking first 5 rows In [3]: df.head() Out[3]: LYLTY\_CARD\_NBR DATE STORE\_NBR TXN\_ID PROD\_NBR PROD\_NAME PROD\_QTY TOT\_SALES PACK\_SIZE **BRAND** Natural Chip 2018-0 2 1000 6.0 NATURAL 1 Compny 175 10-17 SeaSalt175g Red Rock Deli 2018-1 1 1002 58 Chikn&Garlic 1 2.7 150 RRD 09-16 Aioli 150g **Grain Waves** 2019-Sour **GRNWVES** 2 1003 3 3.6 210 1 1 52 Cream&Chives 210G Natural 2019-ChipCo Hony 3 1003 1 106 1 3.0 175 **NATURAL** Soy Chckn175g WW Original 2018-1004 1.9 160 WOOLWORTHS 1 Stacked Chips 11-02 160g Checking data type of variables In [4]: #checking data types df.dtypes Out[4]: LYLTY CARD NBR int64 DATE object STORE NBR int64 TXN ID int64 PROD NBR int64 PROD NAME object PROD QTY int64 TOT SALES float64 PACK SIZE int64 BRAND object LIFESTAGE object PREMIUM\_CUSTOMER object dtype: object Creating Year-Month column Since we are interested in analyzing numbers on monthly basis, therefore, we are extracting year and month from date column and storing it in yyyymm format. #created new column MonthID in yyyymm format from the DATE column In [5]: df["MonthID"] = df["DATE"].str.slice(0,7) df["MonthID"] = df["MonthID"].str.replace(pat="-", repl="") df.head() Out[5]: **BRAND** LYLTY\_CARD\_NBR DATE STORE\_NBR TXN\_ID PROD\_NBR PROD\_NAME PROD\_QTY TOT\_SALES PACK\_SIZE Natural Chip 2018-Compny 0 1000 1 2 6.0 175 NATURAL 10-17 SeaSalt175g Red Rock Deli 2018-RRD SI 1002 1 2.7 58 Chikn&Garlic 150 09-16 Aioli 150a **Grain Waves** Cream&Chives 03-07 210G Natural 2019-ChipCo Hony **NATURAL** 3 1003 106 3.0 175 03-08 Chckn175g WW Original 2018-160 WOOLWORTHS 4 Stacked Chips 1 1.9 **Phase 1: Finding Control Stores Potential Control Store** To find how similar each potential control store is to the trial store, we are going to use two parameters: Monthly Total Sales Monthly Unique Number of Customers Therfore, we are grouping the data by Store number and Months. #group data by STR NBR and MonthID, and taking sum of TOT SALES and unique count of LYLTY CARD NBR. In [6]: measureOverTime = df.groupby(["STORE NBR", "MonthID"]).agg({"TOT SALES": "sum", "LYLTY CARD NBR":pd.Series .nunique}).reset index() #renaming columns to reflect the true meaning of the features measureOverTime.rename(columns={"TOT SALES":"totSales","LYLTY CARD NBR":"nCustomers"},inplace=True) measureOverTime Out[6]: STORE\_NBR MonthID totSales nCustomers 0 201807 206.9 49 201808 176.1 42 201809 278.8 59 3 201810 188.1 44 1 201811 192.6 46 272 201902 395.5 3164 45 3165 272 201903 442.3 50 3166 272 201904 445.1 54 3167 272 201905 314.6 34 312.1 3168 272 201906 34 3169 rows × 4 columns Stores with full observations We are interested in stores, who are operated or had transactions for all 12 months. We will consider only these stores for analysis and leave other stores. For this we simply count MonthID for each store and consider those having equal to 12. #saving store numbers having count of MonthIDs = 12 In [7]: storesWithFullObs = measureOverTime.groupby("STORE NBR")["MonthID"].count().reset index() storesWithFullObs = storesWithFullObs[storesWithFullObs["MonthID"] == 12]["STORE NBR"].unique() storesWithFullObs Out[7]: array([ 5, 7, 9, 14, 1, 2, 3, 4, 6, 8, 10, 12, 13, 17, 18, 19, 20, 21, 22, 23, 24, 16, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 43, 42, 45, 46, 47, 48, 49, 50, 52, 53, 54, 51, 57, 58, 59, 60, 61, 62, 63, 64, 65, 69, 70, 71, 72, 73, 74, 75, 77, 78, 79, 80, 81, 87, 96, 84, 86, 89, 90, 91, 94, 95, 83. 88, 93, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 207, 208, 209, 210, 212, 213, 214, 215, 216, 217, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272], dtype=int64) Changing data type of MonthID from object to integer for ease measureOverTime.dtypes In [8]: Out[8]: STORE NBR int64 MonthID object totSales float64 nCustomers int64 dtype: object In [9]: measureOverTime["MonthID"] = measureOverTime["MonthID"].astype(int) measureOverTime.dtypes Out[9]: STORE NBR int64 MonthID int32 totSales float64 nCustomers int64 dtype: object Data for pre-trial period We know that trail period is from February 2019 to April 2019, and to find how similar each potential control store is to the trial, we will perform the analysis on pre-trial period. Therfore, we are slicing the data for stores having full observations and for time period July 2018 to January 2019. preTrialMeasures = measureOverTime[(measureOverTime["STORE NBR"].isin(storesWithFullObs)) & \ In [10]: (measureOverTime["MonthID"] < 201902)]</pre> preTrialMeasures Out[10]: STORE\_NBR MonthID totSales nCustomers 0 201807 206.9 49 201808 176.1 42 201809 278.8 59 3 201810 188.1 44 201811 192.6 46 3159 201809 304.7 32 272 3160 272 201810 430.6 44 3161 201811 376.2 272 41 3162 272 201812 403.9 47 3163 272 201901 423.0 46 1820 rows × 4 columns **Defining Driving Metrics** We will be using two main metrics for finding the similarity between control strores and trial stores. Correlation Score We will calculate how correlated the performance of each store is to the trial store ■ Correlation value lies between -1 to 1. Value near to 1 means similarity value near to -1 means dis-similarity Absolute Difference Score • We will calculate the absolute difference between the trial store's performance and each control store's performance We will standardize the distance so that the measure ranges from 0 to 1 ■ We will take (1 - standardized distance) so that it has similar effect as correlation score. Value near to 1 means similarity value near to 0 means dis-similarity We will be using these metrics to calculate score for both totSales and nCustomer i.e. we will be having total 4 scores for each control store to the trail store. Then we will combine all these scores to get one score for each store and we sort them in descending order. Store with higest score other than trial store itself becomes the control store of the respective trial store. Let us define the functions to calculate all 4 scores and to store them into a sangle dataframe. It will reduces our time of re-write the code again. In [11]: def corrSales(data, trialStore): It calculates the the correlation score of each store to the trial store obased on total sales. Parameters: data: (dataframe) It takes dataframe contaning store number, months, sales and number o f customers trialStore: (integer) It takes the trial store number from which we need to compare all stores Return: It returns a dataframe with store number and correlation score corrosponding to that store and the trail store df = data[["STORE NBR", "MonthID", "totSales"]] storeNBR = df["STORE NBR"].unique() corr = list() ts = df[df["STORE NBR"] == trialStore].iloc[:,-1] for n in storeNBR: cs = df[df["STORE\_NBR"] == n].iloc[:,-1] corr.append(np.corrcoef(ts,cs)[0,1]) corrSalesDf = pd.DataFrame({"STORE NBR":storeNBR,"corrSales":corr}) return corrSalesDf In [12]: def corrCust(data, trialStore): It calculates the the correlation score of each store to the trial store based on number of custome rs. Parameters: data: (dataframe) It takes dataframe contaning store number, months, sales and number o f customers trialStore: (integer) It takes the trial store number from which we need to compare all stores Return: It returns a dataframe with store number and correlation score corrosponding to that store and the trail store df = data[["STORE NBR", "MonthID", "nCustomers"]] storeNBR = df["STORE NBR"].unique() corr = list() ts = df[df["STORE NBR"] == trialStore].iloc[:,-1] for n in storeNBR: cs = df[df["STORE NBR"] == n].iloc[:,-1] corr.append(np.corrcoef(ts,cs)[0,1]) corrCustDf = pd.DataFrame({"STORE NBR":storeNBR,"corrCust":corr}) return corrCustDf In [13]: def magDistSales(data, trialStore): It calculates the the absoulte distance of each store to the trial store based on monthly sales. Parameters: data: (dataframe) It takes dataframe contaning store number, months, sales and number o f customers trialStore: (integer) It takes the trial store number from which we need to compare all stores Return: It returns a dataframe with store number and absolute distance corrosponding to that store and the trail store 11 11 11 df = data[["STORE NBR", "MonthID", "totSales"]] storeNBR = df["STORE NBR"].unique() dist = list() ts = np.array(df[df["STORE NBR"] == trialStore].iloc[:,-1]) for n in storeNBR: cs = np.array(df[df["STORE\_NBR"] == n].iloc[:,-1]) d = sum(np.abs(ts - cs))dist.append(d) magDistSalesDf = pd.DataFrame({"STORE\_NBR":storeNBR,"distSales":dist}) return magDistSalesDf In [14]: def magDistCust(data, trialStore): It calculates the the absoulte distance of each store to the trial store based on number of custome rs. Parameters: data: (dataframe) It takes dataframe contaning store number, months, sales and number o f customers trialStore: (integer) It takes the trial store number from which we need to compare all stores Return: It returns a dataframe with store number and absolute distance corrosponding to that store and the trail store df = data[["STORE NBR", "MonthID", "nCustomers"]] storeNBR = df["STORE\_NBR"].unique() dist = list() ts = np.array(df[df["STORE\_NBR"] == trialStore].iloc[:,-1]) for n in storeNBR: cs = np.array(df[df["STORE\_NBR"] == n].iloc[:,-1]) d = sum(np.abs(ts - cs))dist.append(d) magDistCustDf = pd.DataFrame({"STORE\_NBR":storeNBR,"distCust":dist}) return magDistCustDf In [15]: def stdMagDis(data, colList): It standarizes the absolute distance so that the measure ranges from 0 to 1. data: (dataframe) It takes dataframe contaning absolute distance score colList: (list) It takes the list of columns which needs to be standardized. To standardize, it uses MinMaxScalar from sckit-learn library Return: It returns the dataframe with (1 - standardized distance) and all other given columns df = data.copy() from sklearn.preprocessing import MinMaxScaler minMaxScalar = MinMaxScaler() df[colList] = minMaxScalar.fit\_transform(df[colList]) for col in colList: df.loc[:, col] = 1 - df.loc[:, col]return df In [16]: def controlStores(data, trialStore): It returns a dataframe with all 4 scores correlation score for sales and number of customers, absol for sales and number of customers of each store to the given trial store. It merge all 4 dataframes returned from different functions into one single dataframe. Parameters: data: (dataframe) It takes dataframe contaning store number, months, sales and number o f customers trailStore: (integer) It takes the trial store number from which we need to compare all stores controlStoresDf = corrSales(data, trialStore) #merging corrSales and corrCust -> controlStoresDf controlStoresDf = controlStoresDf.merge(right=corrCust(data, trialStore), how="inner", on="STORE NB R")  $\#merging\ controlStoresDf\ and\ magDistSales\ ->\ controlStoresDf$ controlStoresDf = controlStoresDf.merge(right=magDistSales(data, trialStore), how="inner", on="STOR") #merging controlStoresDf and magDistCust -> controlStoresDf controlStoresDf = controlStoresDf.merge(right=magDistCust(data, trialStore), how="inner", on="STORE NBR") #Standardising the magnitude distance so that the measure ranges from 0 to 1 controlStoresDf = stdMagDis(controlStoresDf, ["distSales", "distCust"]) #creating a column to show the compared trial store number controlStoresDf.loc[:, "trialStore"] = trialStore #returning final dataframe return controlStoresDf Let's see all four score corrosponding to Trial Store (77) In [17]: controlStores77 = controlStores(preTrialMeasures, 77) controlStores77 Out[17]: STORE\_NBR corrSales corrCust distSales distCust trialStore 1 0.075218 0.322168 0.950903 0.938744 2 -0.263079 -0.572051 0.933198 0.921899 77 1 0.834207 0.317676 0.318530 77 3 4 -0.263300 -0.295639 0.130226 0.157734 77 5 -0.110652 0.370659 0.526858 0.460949 77 255 269 -0.315730 -0.474293 0.418570 0.315467 256 77 0.315430 -0.131259 0.414653 0.333844 77 257 258 271 0.355487 0.019629 0.523884 0.459418 77 0.117622 0.223217 0.877596 0.946401 77 260 rows × 6 columns Let's see all four score corrosponding to Trial Store (86) In [18]: controlStores86 = controlStores(preTrialMeasures, 86) controlStores86 Out[18]: STORE NBR corrSales corrCust distSales distCust trialStore 1 0.445632 0.485831 0.212664 0.433681 2 -0.403835 -0.086161 0.169679 0.366617 1 86 3 -0.261284 -0.353786 0.766059 0.912072 4 -0.039035 -0.169608 0.499755 0.773472 3 86 5 0.235159 -0.253229 0.924119 0.922504 86 268 -0.452182 -0.034273 0.239638 0.414307 255 269 0.697055 -0.098587 0.899881 0.918033 256 86 270 -0.730679 -0.767267 0.836892 0.891207 257 86 271 0.527637 0.267393 0.919228 0.932936 258 86 0.004926 -0.353815 0.438479 0.411326 260 rows × 6 columns Let's see all four score corrosponding to Trial Store (88) In [19]: | controlStores88 = controlStores(preTrialMeasures, 88) controlStores88 Out[19]: STORE\_NBR corrSales corrCust distSales distCust trialStore 0 1 0.813636 0.305334 0.137831 0.340749 88 2 -0.067927 -0.452379 0.109972 0.288056 3 4 -0.745566 -0.361503 0.897187 0.923888 88 0.190330 -0.025320 0.607131 0.731850 88 ... 268 -0.021429 0.672672 0.155314 0.325527 255 88 256 -0.172578 -0.274781 0.706839 0.843091 88 0.710445 0.829040 257 -0.723272 -0.103032 88 271 -0.103037 -0.018831 0.609869 0.733021 258 88 259 272 -0 772772 0 026909 0 284186 0 323185 88 260 rows × 6 columns Let's concatenate scores corrosponding to all three trial stores into a single dataframe In [20]: controlStoresAll = pd.concat([controlStores77,controlStores86,controlStores88],axis=0).reset\_index(drop controlStoresAll Out[20]: STORE\_NBR corrSales corrCust distSales distCust trialStore 0.075218 0 0.322168 0.950903 0.938744 2 -0.263079 -0.572051 0.933198 0.921899 1 77 0.806644 0.834207 0.317676 0.318530 2 77 3 -0.263300 -0.295639 0.130226 0.157734 77 -0.110652 0.370659 0.526858 0.460949 77 0.672672 0.155314 0.325527 775 268 -0.021429 88 776 -0.172578 -0.274781 0.706839 0.843091 88 777 -0.723272 -0.103032 0.710445 0.829040 88 778 271 -0.103037 -0.018831 0.609869 0.733021 88 779 272 -0.772772 0.026909 0.284186 0.323185 88 780 rows × 6 columns Combining all 4 scores to get a single score We'll now combine all 4 scores to get a single score for each store to a trail store. This will be the final score to decide the control store. We will be assigning equal weightage to all four drivers and combine them i.e. we will be taking simple average of all scores. In [21]: #taking average of correlation and absoulte difference difference scores for sales controlStoresAll["scoreNSales"] = (0.5 \* controlStoresAll["corrSales"]) + (0.5 \* controlStoresAll["dist Sales"]) #taking average of correlation and absoulte difference difference scores for number of customers controlStoresAll["scoreNCust"] = (0.5 \* controlStoresAll["corrCust"]) + (0.5 \* controlStoresAll["distCu #we are taking average of sales and number of customers scores controlStoresAll["finalControlScore"] = (0.5 \* controlStoresAll["scoreNSales"]) + (0.5 \* controlStoresA 11["scoreNCust"]) controlStoresAll In [22]: Out[22]: STORE\_NBR corrSales corrCust distSales distCust trialStore scoreNSales scoreNCust finalControlScore 0.571758 0 0.075218 0.322168 0.950903 0.938744 77 0.513060 0.630456 -0.572051 -0.263079 77 0.335060 0.174924 0.254992 0.933198 0.921899 2 0.806644 0.834207 0.317676 0.318530 77 0.562160 0.576369 0.569264 3 -0.263300 -0.295639 0.130226 0.157734 77 -0.066537 -0.068953 -0.067745 77 0.460949 0.208103 -0.110652 0.370659 0.526858 0.415804 0.311954 ... ... 0.672672 0.155314 0.325527 88 0.066942 0.499099 0.283021 775 268 -0.021429 -0.172578 0.706839 0.843091 776 -0.274781 88 0.267131 0.284155 0.275643 -0.723272 -0.103032 0.710445 0.829040 88 777 -0.006413 0.363004 0.178295 778 -0.103037 -0.018831 0.609869 88 0.253416 0.357095 0.305256 88 779 272 -0.772772 0.026909 0.284186 0.323185 -0.2442930.175047 -0.034623 780 rows × 9 columns **Control Store for Trial Store (77)** The store with the Highest score (other than trial store itself) will be selected as the control store since it is most similar to the trial store. In [23]: controlStoresAll[controlStoresAll["trialStore"] == 77].sort values("finalControlScore", ascending=False) Out[23]: STORE NBR corrSales corrCust distSales distCust trialStore scoreNSales scoreNCust finalControlScore 1.000000 1.000000 1.000000 1.000000 1.000000 77 1.000000 77 1.000000 72 0.984567 0.992343 221 0.990358 0.944171 0.991350 0.967761 233 0.903774 0.783232 0.844219 38 0.962834 0.973966 77 0.873033 0.909093 0.891063 15 0.842668 0.747308 0.872450 0.961715 77 0.857559 0.854511 0.856035 77 241 0.577108 0.916208 0.747138 0.925179 254 0.917168 0.934150 0.836159 ... 0.130226 0.157734 -0.067745 3 4 -0.263300 -0.295639 -0.066537 -0.068953 77 -0.585174 130 -0.534877 0.458558 0.358346 77 -0.088266 -0.075787 -0.063308 235 247 -0.631050 -0.621034 0.498955 0.402757 77 -0.066047 -0.109139 -0.087593 -0.550834 -0.652527 77 -0.043972 -0.140199 -0.092086 95 0.462890 0.372129 77 71 -0.806751 -0.590735 0.266465 0.307810 -0.270143 -0.205803 -0.141463 260 rows × 9 columns Store 233 is the Control Store corrosponding to Trial Store 77 because it has the highest score among all other stores **Control Store for Trial Store (86)** The store with the Highest score (other than trial store itself) will be selected as the control store since it is most similar to the trial store. In [24]: controlStoresAll[controlStoresAll["trialStore"] == 86].sort values("finalControlScore", ascending=False) Out[24]: corrCust distSales distCust trialStore scoreNSales scoreNCust finalControlScore STORE\_NBR corrSales 1.000000 1.000000 340 1.000000 1.000000 1.000000 1.000000 1.000000 407 0.877882 86 0.920039 0.942013 155 0.942876 0.962197 0.985097 0.963986 362 0.770778 0.963195 0.965723 0.871999 0.734415 86 0.827175 0.894883 0.861029 367 0.855339 0.919935 0.934426 0.841833 0.840085 390 0.759864 0.749701 0.923802 0.926975 0.838338 -0.840413 -0.697245 0.329577 0.327869 86 -0.255418 -0.184688 -0.220053 361 308 52 -0.601629 -0.594459 0.024254 0.020864 -0.288688 -0.286798 -0.287743 86 -0.872693 -0.815097 0.160064 0.357675 -0.356315 -0.228711 -0.292513 372 299 42 -0.745720 -0.664952 0.009233 0.019374 86 -0.368243 -0.322789 -0.345516 146 -0.775127 -0.654598 0.008268 0.005961 86 -0.353874 -0.383430 -0.324319 260 rows × 9 columns Store 155 is the Control Store corrosponding to Trial Store 86 because it has the highest score among all other stores **Control Store for Trial Store (88)** The store with the Highest (other than trial store itself) score will be selected as the control store since it is most similar to the trial store. In [25]: controlStoresAll[controlStoresAll["trialStore"] == 88].sort\_values("finalControlScore", ascending=False) Out[25]: STORE\_NBR corrSales corrCust distSales distCust trialStore scoreNSales scoreNCust finalControlScore 602 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 745 237 0.308479 0.947326 0.955903 0.987119 88 0.632191 0.967223 0.799707 690 0.731857 0.939466 0.693901 0.818501 88 0.712879 0.878984 0.795931 585 0.450029 0.815792 0.706057 0.867681 88 0.578043 0.841737 0.709890 69 626 113 0.495763 0.862632 0.680855 0.781030 88 0.588309 0.821831 0.705070 604 -0.662427 -0.503735 0.175513 0.310304 -0.243457 -0.096715 -0.170086 0.087493 -0.825850 0.008248 0.001171 88 765 0.047871 -0.412339 -0.182234 653 -0.690590 -0.547399 0.195038 0.286885 88 -0.247776 -0.189016 -0.130257 747 -0.642329 -0.660672 0.247620 0.291569 88 -0.197355 -0.184552 -0.190953 559 -0.519338 -0.282575 0.005984 0.015222 88 -0.256677 -0.133676 -0.195177 260 rows × 9 columns Store 237 is the Control Store corrosponding to Trial Store 88 because it has the highest score among all other stores **Assessment of Control Stores prior Trial Period** (July 2018 - January 2019) Now that we have found a control store, let's check visually if the drivers are indeed similar in the period before the trial for control store, trial store and other stores(taking all other stores together). Let us define a function to reduce our efforts In [26]: def controlStoresPriorTrialPeriodViz(data,controlStore,trialStore,savfig=False): It create two line plots, one for sales and another for number of customers for the given control a nd trial store to check visually if the drivers are indeed similar in the period before the trial. Parameters: data: (dataframe) It takes dataframe with store number, months, sales and number of cus tomers for all stores controlStore: (integer) It takes control store number trialStore: (integer) It takes corrosponding trial store number savfig: (boolean) It asks for saving the figure on the local system or not, by default it is False. It creates a single line for stores other than control and trial store by aggregating values by mea n 11 11 11 df = data.copy() df["storeType"] = df["STORE\_NBR"].apply(lambda x: "Control" if x==controlStore else "Others") df.loc[df["STORE NBR"]==trialStore,"storeType"] = "Trial" df = df[df["MonthID"] < 201902]df["MonthID"] = df["MonthID"].astype(str) plt.figure(figsize=(15,4)) plt.subplot(1,2,1)sns.lineplot(data=df, x="MonthID", y="totSales", hue="storeType", estimator="mean", err\_style=None, style="storeType", \ dashes=[(2,1),(1,0),(5,1)])plt.xlabel("Time Period (yyyymm)", size=13) plt.ylabel("Sales", size=13) plt.subplot(1,2,2)sns.lineplot(data=df, x="MonthID", y="nCustomers", hue="storeType", estimator="mean", err style=Non e, style="storeType", \ dashes=[(2,1),(1,0),(5,1)]) plt.xlabel("Time Period (yyyymm)", size=13) plt.ylabel("Number od Customers", size=13) plt.suptitle(f"Comparing stores performance (prior trial period) \nTrial Store: {trialStore} Contro 1 Store: {controlStore}") if savfig==True: path="C:/Users/amitm/Jupyter Notebools/Quantium/Task 2/"+"Stores Performance ("+str(trialStore) +").jpeg" plt.savefig(path,dpi=500) Visualizing stores performance in the period before the trial for Control Store (233) and Trial Store (77)

<ul> <li>We can from tria</li> <li>This sho chooser</li> <li>Note: Measo</li> <li>Let's visu (86)</li> </ul>	ualise stores	s perfoi	rmance in t	the perio	d before th				e (155) a	nd Tria
1000 900 800 700	oresPriorTri	alPerio		paring stores	performance ( : 86 Control S	(prior trial p		store=86  store Other Trial Contr	Гуре	=True)
from tria  This sho	see that lines a	re almost osen cont	d (yyyymm) over lapping ea trol store (155)	is correct ar	both charts fo	or control st	ore and tria	ne Period () Il store and	yyyymm) d others sto	
(88)	oresPriorTri		odViz(measur	reOverTime	performance (: 88 Control S	Store=23°	7, trials			
from tria  This sho	see that lines a	re almost	d (yyyymm) over lapping ea		both charts fo	or control st	ore and tria	ne Period () Il store and	yyymm) d others sto	
Pl (February Our goal is t and make co	nase 2: A y 2019 - Apr o see if there ha	<b>\SSES</b> il <b>2020)</b> s been an	ssment	<b>of Tria</b>	in trail stores.	We now h	ave control	stores to	compare co	ompare t
Assess Calculati	ment of s	ales o	of Trial St for Control	Store (77) $= \frac{\sum Tric}{\sum Contr$	Sale al Store's So rol Store's S	$ales\ (prions Sales\ (prions Sales$	r trial per or trial pe	$rac{iod)}{riod)}$		
reTrialMe scalingFa 1.0236173 Creating	ctorForContr 03289553 <b>a dataframe</b> trolSales =	olSales with c	ontrol stor	re's data d	reTrialMeas	sures ["S?	CORE_NBR'	']==233]	["totSal	es"].s
	233 201808 233 201808 233 201808 233 201810 233 201812 233 201812 233 201902 233 201902	7 290.7 3 285.9 228.6 0 185.7 1 211.6 2 279.8 1 177.5	7							
	233 201900 233 201900 233 201900 233 201900 control store  trolSales ["controlSales   "controlSales   "con	1 158.6 5 344.4 6 221.0 e's sale	s controlSales						orForCon	trolSa
2699 2700 2701 2702 2703 2704 2705 2706	233 201803 233 201803 233 201803 233 201810 233 201812 233 201903 233 201903 233 201903	7 290.7 8 285.9 9 228.6 1 185.7 1 211.6 2 279.8 1 177.5 2 244.0	297.565550 292.652187 233.998916 190.085733 216.597421 286.408121 181.692071 249.762622							
trialSale	233 201904 233 201906 233 201906  trial store's  s = measureC  trolSales = trolSales.re	sales in	352.533799 226.219424  The same [measureOve	e datafram erTime["S] s.merge(tr	<pre>FORE_NBR"] =</pre>	on="Month	nID")			lace= <b>1</b>
STORE_  0 1 2 3 4 5	233 201807 233 201808 233 201809 233 201810 233 201811 233 201812 233 201901	290.7 285.9 228.6 185.7 211.6 279.8 177.5	controlSales 1 297.565550 292.652187 233.998916 190.085733 216.597421 286.408121 181.692071	296.8 255.5 225.2 204.5 245.3 267.3 204.4						
Now that we	233 201902 233 201903 233 201904 233 201905 233 201906   The properties of the prope	le sales fi	igures for the c	control store,	we can calcu	late the pe	rcentage di	fference b	etween the	escaled
"trialSal scaledCon STORE  0 1 2 3	NBR MonthID  233 201807  233 201808  233 201809  233 201810	290.7 285.9 228.6 185.7	controlSales 1 297.565550 292.652187 233.998916 190.085733	trialSales pe 296.8 255.5 225.2 204.5	ercentageDiff			Sales"]	- scaled	Contro
4 5 6 7 8 9 10 11	233 201811 233 201812 233 201901 233 201902 233 201903 233 201904 233 201905 233 201906	211.6 279.8 177.5 244.0 199.1 158.6 344.4 221.0	216.597421 286.408121 181.692071 249.762622 203.802205 162.345704 352.533799 226.219424 gnificant!	245.3 267.3 204.4 235.0 278.5 263.5 299.3 264.7	0.132516 0.066716 0.124980 0.059107 0.366521 0.623080 0.151003 0.170103					
We need to a Null Hypoth Alternate Hy Let's calcul stdDev = stdDev	the difference check and confinences: There is no spothesis: There is no spothesis: There is no scaledControl of the control o	m this sta o differen e is a diffe	etistically, and vote in sales in the sales on the based on the sales	rial and pre-ts in trial and he scaled p	trial period pre-trial period ercentage di	d fference in	the pre-tr	ial period	ff"].std	1 ()
Let's calcul  Our focus will  We are testing	ate the t-value  ill be for trial per  ng with a null hy  trolSales["t	pothesis o	t-va	$alue = \frac{Per}{}$	$rac{rcentageDi}{StandardI}$	fference Deviation				racted z
0 1 2 3 4 5 6 7	NBR         MonthID           233         201807           233         201808           233         201809           233         201810           233         201811           233         201812           233         201901           233         201902           233         201903           233         201904	290.7 285.9 228.6 185.7 211.6 279.8 177.5 244.0 199.1	297.565550 292.652187 233.998916 190.085733 216.597421 286.408121 181.692071 249.762622 203.802205	296.8 255.5 225.2 204.5 245.3 267.3 204.4 235.0 278.5	0.002573 0.126950 0.037602 0.075830 0.132516 0.066716 0.124980 0.059107 0.366521	1.335911 2.502571 1.183534 7.339116				
9 10 11 The 95th persignificant. Since there is the Hence, degrees of the Hence of	233 201904 233 201905 233 201906  Procentile of the series of freedom  The end because with the series of the seri	158.6 344.4 221.0 t-distributher pre-tries = 6 (7 - 1)	162.345704 352.533799 226.219424 tion with the a	263.5 299.3 264.7 appropriate	0.623080 0.151003 0.170103 <b>degrees of fi</b>	12.476373 3.023650 3.406093			nypothesis	s is stat
degreesOf stats.t.p 1.9431802  Comparing scaledCon STORE_I 7 8	Freedom = 6 pf(0.95,df=d) 803927816  t-value and 95t trolSales[sc  NBR MonthID t 233 201902 233 201903	h percent caledCont otSales caled.0	tile of the t-distrolSales ['controlSales tr 249.762622 203.802205	stribution for "MonthID" ] rialSales per 235.0 278.5	or trial period  isin([20]  centageDiff  0.059107  0.366521	tValue 1.183534 7.339116				
9 <b>Observat</b> We can obse	233 201904	158.6 alue is mu	162.345704 ch larger than	263.5 the 95th per	0.623080 1 centile value o	2.476373 of the t-dist	ribution for	March and	d April - i.e.	the inc
Let's create	trial store in Mar ng the Sales a more visual ver rol store's sales.	ersion of th	ıl and Cont			ere, trial sto	re, other sto	ores and th	ne 5% and	95% co
Let's create level of control Let's define d	a more visual verol store's sales.  a function for this  olstoreViz (do  ots a line property and 95% on tile confidence on the contract of the confidence on the contract of the confidence on the contract of the confidence on the confidence on the contract of the confidence on the confidence on the confidence on the confidence on the confidence of the c	swhich we lata, con lata,	all and Continus by plotting to the can use for a trolstore, to the continue of the continue o	the sales of the sales of the sales of the stores of custom trol store akes datard at takes the state only two masks to sales of the store the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only two masks to sales of the store that takes the state only the store that takes the state only the store that takes the state of the store that takes the state of the state of the state of the store that takes the state of the sta	che control storalso  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control strial storal  ave the cha  chan control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  control  cave the chan  cave the cha	ail Store  number of  store in  store's in  e's number  iation whetsales",  art on the  l and tri  fial"  ID", on, "s  ] * (1 +  Store]["More interval  ] * (1 -  ence interval  ask 2/"+  tore=77,  eriod	=False):  e, Control of custor  amber, monumber, monumber er nich is a "nCustor ne local ial store 2 * stdl STORE_NBE nthID"] 1"] * ler  2 * stdl erval"  amean(). e", size=1 colStore)  amean+" All stdDev=5	ol Store mers ove  onths, s  used in mers" to  system  e by agg  crolStor  e"]) Dev) R"] n(df[df[ Dev)  "", size=	, and Ote r the wh ales and hypothes plot us ot not, regating e else " "STORE_N  20 15)	tore)+
Let's create level of contribute of contribu	a more visual verol store's sales.  a function for this  olstoreViz(conts a line program and 95% ntile confidence conting and 95% ntile confidence contraction and contraction	swhich we lata, con lata,	and Continus by plotting to the can use for a strolstore, the can use for a strolstore, the can use for a strolstore, the can use for a strong (integer) in the case (integer) i	the sales of the stores of customs of customs trois to reconstruction of customs are detailed as the state of	the control store  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control st  trial store  ave the cha  han control  ave the cha  han control  control store  ave "Cor  pe"] = "Tr:  BR", "Month:  Store][on:  g=controls  confidence  Istore][on:  period", con  frame with  confidence  avalues "tore  controls  confidence  store Type  (3,1)])  Quantium/Ta  33, trials  over the p  e: 77 Control  avalues "tore  con  con  con  con  con  con  con  co	ail Store  number of  store in  store's in e's number intion with the sales", art on the  land tri  ntrol" in ial"  ID", on, "s  core]["More interval  ore, est:  n().max() "skyblue"  pe", est:  n().max() "skyblue"  core]["More interval  ore, est:  ask 2/"+i  tore=77,  eriod  I Store: 233	=False): e, Control of custor  imber, monumber, monumber erich is in "nCustor ine local ital store 2 * stdI STORE_NBE ithID"] 1"] * ler  2 * stdI erval"  amean(). e", size=1 colStore)  hame+" All stdDev=s	eners over  onths, s  used in mers" to system  e by agg  crolStor  e"]) Dev) R"] n(df[df[ Dev)  "", size=	, and Other the who ales and hypothes plot us ot not, regating e else "  "STORE_N  "STORE_N  con="totS  con="totS	chers Sole ty  I number  Sistes  Sing co  by def  value  There  T
Let's create level of control tet's define and tet's define and to mers for the second and to mers for the second and to mers for the second and the second	a more visual verol store's sales.  a function for thi  olstoreViz (do  ots a line product of the confidence of the conf	ersion of the swhich we have a lata, considered and a lata, consider	all and Continus by plotting to the process of the period of of the p	the sales of the stores of customes of cus	che control storalso  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control st  trial storal  andard dev.  values "toral  ave the cha  chan control  control storal  control sto	ail Store number of store in store in store is e's numbe iation wi tSales", art on the land tri land t	=False): e, Control of custor umber, monumber enumber enich is to "nCustor the local dal store at x==cont storeType 2 * stdI STORE_NBE thID"] L"] * len  2 * stdI erval"  Trial P  stdDev=s  stdDev=s  stdDev=s	eriod	, and Other the who ales and hypothes plot us ot not, regating e else "  "STORE_N  "STORE_N  con="tots"	chers Sole ty  I number  Sis tes  Sing co  by def  value  Others  TBR"]==
Let's create level of control Let's define in the series of the series o	a more visual verol store's sales.  a more visual verol store's sales.  a function for this  olstoreViz (do  ots a line proper sand 95% and 95% and 95% and 95% and 95% and 95% control solosolosolosolosolosolosolosolosolos	ersion of the swhich we have a construction of the same this, it will be the same this the same th	all and Continis by plotting to this by plotting to the plotting to the can use for a stroistore, the color of the can use for a stroistore, the can use for a stroistore, the can use for a stroistore for a str	the sales of the stores of the stores of customs trois stores of customs trois stores of customs trois stores of customs trois stores of customs of custom	control storals  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control storal storal  andard dev:  values "tor  andard dev:  values "tor  and control  ave the cha  chan control  control storal  control storal	ail Store  number of  store in  land tri  land tri  land tri  store] ["s  store] ["s  pe", est:  a().max()  "skyblue"  store] ["s  pe", est:  a().max()  store] ["s  pe", est:  a().max(	False):  e, Control of custor of cus	eriod  col Store  ners ove  nonths, s  ised in  ners" to  system  by agg  crolStor  cr	, and Other the whole ales and hypothess plot us ot not, regating e else "  "STORE Not trials and trials on store points and trials	t also  anumber  tore)  ales",
Let's create level of continuation of the level of the lev	a more visual verol store's sales.  a more visual verol store's sales.  a function for thi  olstore viz (do  ots a line provide a line provid	ersion of the swhich we have a considered interest of cust and interest	al and Continis by plotting to the process of the continuous of the continuous plants of the con	the sales of the stores as trial stores of custome trol stores as trial stores as trial stores as the stores are	the control storals  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control st  trial storal  avalues "toral  ave the characterial storal  frame control  fram	store in the store in the store in the store in the scale	False):  e, Control of custor  amber, mo number er nich is n "nCustor ne local  ial store 2 * stdl STORE_NBE nthID"]  """ * len 2 * stdl erval"  amean(). e", size=1 colstore)  amean(). e	eriod  col Store mers ove  conths, s  ased in mers" to system  by agg  crolStor  crolS	, and Ott r the wh ales and hypothes plot us ot not, regating e else " "STORE N  ales and one and in  ales and in	t also there
Let's create level of control  Let's define of the control  To define of the control  Let's define of the control  Let's define of the control  Let's define of the control  Control  Control  Control  Control  True)  Control  Con	a more visual verol store's sales.  a function for thi  olstoreViz (do  ots a line production of the product of	swhich was attacked and a control of the same of the s	all and Continis by plotting to the plotting to the can use for a strong to the can us	the sales of the stores and the stores are s	the control sto  also  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control st  trial store  andard dev:  values "to:  ave the cha  than control  control store  control store  control store  confidence  as confidence	store in the store	False):  Fal	pol Store mers ove mers ove onths, s  used in mers" to system e by agg crolStor  e"]) n(df[df[ Dev)  n(df[df[ Dev)  n(df[df[ Dev)  nothin, s  onthin, s  contine  nonthin, s  contine  nonthin, s  contine  nonthin, s  contine  nonthin, s	, and of rether who ales and hypothes plot us of not, regating e else "  "STORE No ales and in ales an	t also  all numbers  all number
Let's create level of control to the steam of the state o	a more visual vertical vertica	swhich w  swhich w  lata, con  la	and Continis by plotting to this by plotting to the can use for a stroit of the can us	the sales of the stores of customs of custom	the control store  also  e, stdDev, or  mers of Tra  e's sales/  frame with  s control st  randard devia  values "tor  ave the cha  than control  coda x: "Cor  coe"] = "Tra  axi, "Month;  control st  control st	n, savfigable in the store in t	False):  Fal	ol Store  ol Store  ners ove  onths, s  used in  mers" to  system  e by agg  crolStor  e'"])  ol (df[df[  olev)  nean", e  el (0.5)  min() -  ob, onths, s  contine  clotdev,  onths, s  conths, s  conthro  ners" to  onths, s  conthro  ners" to  onths, s  conthro  ners" to	rathe what ales and hypothes plot us of not, regating e else " "STORE_N" "STORE_N" "STORE_N" "on="tots" ales and in ales and i	t also  anumber  also  anumber  also  anumber  also  anumber  ales",  anumber  ales",  anumber  anumbe
Let's create level of control  Let's define.  Let's define.  def contr  With 5 perced.  Param  tomers for  ing measur  is False  if control  Control  Control  Control  Totol  To	a more visual verol store's sales.  a more visual verol store's sales.  a function for thi  olstoreViz (cots a line property of the sale o	swhich was which was also as a series of the same of t	and Continated in the process of the period	the sales of the sales of the sales of the stores of customs of cu	che control storals  also  als	n, savfig- ail Store number of store in store in store's in store's in e's number is store's in e's number is art on th stales", art on th store] ["s store] ["s store] ["s store] ["s in () .max() "skyblue" in () .max() "store in th three trial store in e's number in () .max() "store in th three trial "store in e's number in () .max() "s in () .m	False):  Processor  Interpolation  I	onths, s  ised in mers ver  onths, s  ised in mers to system  by agg  crolStor  in (df[df[  oev)  in (df[df[  oev)  onths, s  contine  is testi  onths, s  conthro mers to	rathe who ales and hypothes plot us of not, regating e else " "STORE N "STO	chers sole ty linumber list tes ling co by def value  Others  BR"]=  Hone,  ales",  al
Let's define def control  It pl  Ing with 5  As ease  It cr  Ing measure  It pl  Ing measure  It pl  Ing measure  It cr  Ing measure  Ing	a the Sales a more visual verol store's ales a function for thi  olstoreViz (cots a line property and store) a function for thi  cots a line property and store a line propert	sersion of the sersio	al and Contains by plotting the set of contains and set of contains the set of contains and set of contain	the sales of the sales of the stores of customers of cust	the control store  also  also  a, stdDev, or  a, stdDev, or  a, stdDev, or  are sales/  frame with  s control store  andard dev,  values "tor  ave the char  than contro.  ave the char  than control  ave the char  control store  control stor	## A C. 18335.  ## A C. 18335.	False):  Fal	pl Store mers ove mers ove nonths, s  sed in mers" to system syst	rand of the whole ales and hypothes plot us of not, regating e else "  "STORE N  "STORE N  "I ales and in ales and	t also thore to the serior of
Let's define.  Let's define.  def contr  It pl  ng with 5  perce  d.  Param  tomers for  ing measu  is False  It cr  n """  df =  df ["s  df.lo  Contr  Control  if on  cut  cut  cut  cut  cut  cut  cut  cu	regithe Sales a more visual verol store's sales. a function for this olstoreViz (d ots a line product of the control of the co	strion of the swhich we have the same t	al and Contents of the content	the sales of the sales of the sales of the stores are sales of the stores are sales at the sales	the control sto  the control sto  also  al	## A CONTROL OF THE PROCESS OF THE P	False):  Fal	stomers.  stomers.  stomers.  stomers.  stomers.  stomers.  stomers.  stor thro  ners" to  stomers.  stor thro  ners" to  stores.  stores.	rand of the whole ales and hypothes plot us of not, regating e else "  "STORE N  "STORE N  "I ales and in ales and	t also  all number  is tes  by def  value  tothers  BR"]==  None,  in umber
Let's create Let's define Let's define Let's define  def contr  If pl  g with 55  d.  Param  tomers for  ing measu  is False  if cr  n  """  df =  df ["s  df ["s  df ["s  df ["s  df ["s  df ["m  contr  contr  contr  contr  contr  contr  contr  contr  contr  df =  df ["M  name=  if on  else:  "storeTyp  plt.x  plt.s  if sa  plt.f  ytext  plt.t  plt.s  if sa  Param  tomers for  controls  for  controls  ing measu  tomers  controls  trial  standard  plt.s  controls  for  controls  for  controls  for  controls  ing measu  tomers  controls  for  controls  for  controls  for  controls  for  control  ing measu  tomers  for  control  ing measu  ing tomers  ing	re totsales  andre visual vero  rol store's sales.  a function for thi  clistore'viz (cl  cots a line re  s' and 95s  ntile confid  re saver  data:  rall month  contri  stadbe  on:  re saver  data:  ("STORE  class   God  class	rsion of the serion of the ser	all and Continis by plotting to this by plotting to the by plotting to the by plotting to the by plotting to the continuation of the by plotting to the by plotting t	trial Store  akes data  akes data	the control store  also	and it will is three trial is to re in the t	### False :  ### False :  ### Control  ### C	and store  and in mers to system  by agg  crolStor  and [df [df [  and [  and [df [  and	, and of r the wh ales and hypothes ot not, regating e else " "STORE_N "STORE_N "TILLES on="tots on"	chers sole ty in number is tes ing co by def to value  Chers BR"] ==  None,  ales",  ales",  "noust  chers ales ales ales ales ales ales ales ale
Let's create leter of control of a control o	and the Sales a more visual verol store's sales. a function for thi olstore's sales. a function for thi contain trial stadde on: for saver deates a sing data ["STORE deates a sing data ["STORE conterpyel"] = claf ["STORE c	ersion of the er	al and Continis by plotting to this by plotting to the can use for a second and the can use for a secon	the sales of the s	the control sto  also  a	## A CANTER STORE AND A CANTER S	False):  False):  False):  False):  Fy Control  Indianation  Fy Store  Fy St	and store	ng and of r the wh ales and hypothes plot us ot not, regating e else " "STORE_N "STORE_N "TORE_N "TORE_N "TORE N "TORE	chers Society  cole ty  cole t
Calculati  Standard  Stores  Control  C	and the Sales a more visual verols afore is all verols ales. a function for this control state is all month control a sales all month control a sale	resion of the serion of the se	al and Continis by plotting to this by plotting to the can use for a can	the sales of the sales of the sales of the stores of customes of c	the control sto  also  a	## A CONTROL OF THE PROCESS OF THE P	Fralse):  Fralse):  Fralse):  Fralse):  Francontrol  Fran	of Store  of Store  mers ove  onths, s  defines  on	, and of r the wh ales and hypothes of not, regating e else " "STORE N "STORE N "STORE N "Trestyle 20 15)  r(trials on="tots on="	chers is a sole ty and the sole ty along to the sole ty along the sole ty alon
Let's define  Let's define  Let's define  Control  Let's define  Control  Let's define  Let's define	ang the Sales  and the visual service of the sale of t	distribution of the service of customer of the service of the serv	al and Continate of the continuation of the co	the sales of the sales of the sales of the sales of the stores is other stores of customers of the sales of t	the control sto  also  a	## A CALLED TO STORY	False):  Fal	and store	, and of r the wh ales and hypothes plot us of not, regating e else " "STORE_N "STORE_N "Talsor rr_style  20 15)  r(trials on="tots on="tots on="tots on="tots on="tots on" ales and ugh which plot us rialstore p	chers so a control of the solution of the solu
Let's create Let's create Let's define.  Let's define.  Let's define.  def control  ing measur  is Faise  It cr  n """  df =  df("so  df("so  df("so  control  control  control  control  ing measur  tomes  if on  else:  "storeTyp  plt.x  plt.x  plt.x  plt.s  plt.s  for on  else:  controlst  This  standard  is ca  controlst  True)  Calculati  scalingFa  scalingFa  controlst  True  Calculati  scalingFa  controlst  True  Calculati  scalingFa  controlst  True  Calculati  scalingFa  controlst  True  Calculati  scalingFa  control  ing measur  tomers for  control  ing measur  tomers for  control  ing measur  tomers  control  ing measur  tomers  for  control  ing measur  tomers  control  ing measur  tomers  for  control  scale  control  control  scale  control  control  scale  control	ang the Sales ang the Sales and are visual verices alone of the sales and are visual very and a sales and are visual very and are visual v	estation of the estate of the	al and Continate in the product of the can use for a control continuation of the control contr	the sales of the s	the control store  also	## A 18335 ## A 19 A 18335 ## A 19 A 18335 ## A 19 A 18335 ## A 18	False):  Fal	eriod  er	, and of , and of , and of r the wh ales and hypothese ales and hypothese of else " "STORE N	eners s  there is  there i
Let's create Let's define.  Let's create  Let's cre	ang the Sales ang the Sales and are visual verices alone of the sales and are visual very and a sales and are visual very and are visual v	estation of the estate of the	and Contines by plotting the can use for a c	the sales of the s	to to its control  frame with  frame control  frame with  frame control  frame with  frame	## A CONTROL OF THE PROCESS OF THE P	False):  Fal	eriod  er	, and of , and of , and of r the wh ales and hypothese ales and hypothese of else " "STORE N	ners s  number  number
Let's create level of content let's define let's define let's define let's define def content grith part grith	### Sales	resion of the serion of the se	and Continuation of the control of t	the sales of the s	to the control sto  also  also	## A CANDERS  ##	False):  Fal	polistore	my and i  ales and hypothesis of not, regating e else "STORE N "ST	e. the indicate of the indicat
Control  Con	### Sales	resion of the serion of the se	and Contents  an	the sales of the s	the control store in the contr		are trial perimonths.  Trial P  Trial P	ol Store  ol Store  ol Store  onths, s  sed in  ners 'to  osystem  se by agg  crolstor  sel (aff)  crolstor  crolsto	my and i  ales and hypothesis of not, regating e else "STORE N "ST	e. the in a series of the individual of the indi
### Control  ### C	### Sales	distribution of the same of th	and Control  as political  as	trials and a series of a serie	the control store of the contr	## A	False):  Fal	or the state of time  or this, s  see in this, s  set in this,	ny and of  r the wh  ales and hypothesis  by and hypothesis  control  regating  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  restricts  ales and hypothesis  restricts  re	t also to the same of the same
Control  def contr	## Sales  ## Sales  ## A more visual vertication for the control  ##	distribution of the same of th	and Control  as political  as	the sales of the s	the control store of the contr	## Control of the con	stribution for customer and the customer	and store of the s	ny and of  r the wh  ales and hypothesis  by and hypothesis  control  regating  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  ales and hypothesis  restricts  restricts  ales and hypothesis  restricts  re	tore)  tore  from  ales  from
Let's create  Le	### Sales	total and	al and Cont  is by plotting to  is by plotting to  is can use for or  is significate  is can use for or  is	the sales of the s	### ### #### #### ####################	## A	PFAISE):  PFAISE): PFAISE):  PFAISE): PFAI	and store of the s	ny and of r the wh ales and hypothes plot us of not, regating e else " "STORE N "STO	tores  is state  in umbe  is tes  in umbe  in
Cobservat  Let's create level of cont  Control  Let's create level of cont  Let's create level of cont  Control  Let's create level of cont  Let's create level of cont  Control  Let's create level of cont  Let's	### ### ### ### ### ### ### ### ### ##	side interest interes	and Control  and by plotting to the control  and control	tomers  tomers	the control store  the control store  the control store  the control store  also  al	## 11 - ## 12	and the period of customers an	and store of the store of time	mand of the who ales and hypothesis of the store protects on and and and and and and and and and an	chers  ch
## ASSESS	### Sales	strict of the service	al and Continity  is by plotting to the property of the proper	trialsales of the sales of the	the control store  the control store  the control store  the control store  also  al	# 10	False):  Fal	or through a store, of the store of the stor	mand of the state	chers since ty control of the

Discontinue  The control of the cont	Assessment of Number of Dustions and Country of State of	100 san										
Accessment of Sales of Trial Store (83)  Containing Sales of Trial Sto	Assessment of Select of Trial Store (88)  Classification of Select o	Observa The results	Control 95th % contro	nfidence inter 3 201800	9 201810	Time	Period (yyyy) to its contro	ymm) ol store in t	902 20°	1903 20		201 ormano
Controlled	## 18 Part   Par	Assess Calculat	sment of Sing Scaling	Sales (	of Trial St	control store in core (88)  Store's Sa	all three tri	ial months	]["totS	ales"].s	sum()/ preTr	ialMe
March   Marc	1	1.0015583 scaledCod scaledCod STORE	3306649594  ntrolSales, ntrolSales  E_NBR MonthID  237 201807  237 201808	totSales 1448.4 1367.8	controlSales 1 1450.657086 1369.931485	trialSales pero 1310.00 1323.80	entageDiff  0.096961  0.033674	tValue  2.897136 1.006168				
September 1997 -	proposed in the control of the contr	4 5 6 7 8 9	237 201811 237 201812 237 201901 237 201902 237 201903 237 201904 237 201905	1397.6 1265.0 1219.7 1404.8 1208.2 1204.6 1199.3	1399.777923 1266.971288 1221.600696 1406.989143 1210.082775 1206.477165 1201.168906	1382.80 1325.20 1266.40 1370.20 1477.20 1439.40 1308.25	0.012129 0.045959 0.036673 0.026147 0.220743 0.193060 0.089147	0.362408 1.373227 1.095756 0.781270 6.595668 5.768527 2.663672				
### SPECIAL STATES AND AND AND CONTROL OF THE AND C	Expectations of the control of the c	significant Since there Hence, deg  degrees0 stats.t.; 1.894578	are 7 months in grees of freedom  fFreedom = 7  ppf(0.95,df=	n the pre-tr	rial period ) Offreedom)				check w	hether the	e hypothesis is	statis
Visualizing the Sales of Trial and Cerebral Provision in the principle release the remain the standards referenced to the type principle release the remain the standards referenced to the standards remain the remain the standards remainded to the standards remaind	Assessment of Number of Customers of Trial Store (8)  Characteristics and State of Trial and Control Store  Sales over the period  Sales	STORE_7 8 9	NBR MonthID  237 201902  237 201903  237 201904	totSales 1404.8 1208.2	ntrolSales ["  controlSales tr  1406.989143  1210.082775	'MonthID"].  ialSales perce  1370.2  1477.2	isin([201 entageDiff 0.026147 0 0.220743 6	tValue .781270	903,201	904])]		
The Board Course Board 2019  The Course Board	This Store of Coronal Box 227	We can obsin the trial s  Visualizi  Let's create level of concontrols	serve that the t-vectore in March and the Sale a more visual vectors sales	nd April is  s of Tria  version of the	statistically great  al and Cont  this by plotting t	rol Store he sales of the	e control stor	re, trial sto	res, other	<sup>-</sup> stores and	d the 5% and 9	5% cor
Concerning Control of the control state of the destination for the control of the	Control State And Annual Control State State Supplemental State of the Control State	1400 1200 1000 800								Period		
Assessment of Number of Customers of Trial Stora (88)  Calculating Scaling Factor for Control Stora's Number of Customers  ***********************************	Assessment of Number of Customers of Trial Store (88)  Calculating Scaling Factor for Control Store's Number of Customers  Calculating Scaling Factor for Control Store's Number of Customers  Calculating Scaling Factor for Control Store's Number of Customers  Calculating Scaling Factor for Control Store's Number of Customers  Calculating Scaling Factor for Control Store's Number of Customers  Calculating Scaling Factor for Control Store's Customers  Calculating Scaling Factor for Customers  Calculating Scaling Factor for Customers  Calculating Scaling Factor for Customers  Calculating Factor for Customers  Calculati	200 0 Observa	Others Trial Control Control 95th % of 201807  201807  201807  201807	onfidence inte 08 20180 rial in store	erval  201810  288 is significa	Time	Period (yyy	rymm) ol store in t	he trial pe	eriod as the		201 ormano
### Color Process of the Color	TOTAL PLANE AND CONTROL OF THE PROPERTY OF THE	Assess CalculatescalingFixe()	sment of I	Numbe Factor	er of Cust	comers o Store's Nu	f Trial S umber of "STORE_NE	Store (	88) ners	tomers"]		
4 927 20101 192 104 105 105 20450 193 104 105 20450 193 104 105 20450 193 204 105 204	297 A2981   122   124	0.9943502 scaledCod scaledCod STORE 0 1	2824858758  ntrolCust, s ntrolCust  E_NBR MonthID  237 201807  237 201808  237 201809	nCustom	ners controlCus 128 127.276836 135 134.237288 126 125.288136	t trialCust properties 129 131 131 134	ercentageDiff 0.013539 0.024116 0.010281	tValu 1.82702 3.25443 1.38745	e 6 5			
Significant.  Significant Transition in the pre-sist period  Hence, degrees of freedom = 5 (7 - 1)  **Significant Transition = 5 (8 - 1)  **Signif	ingredient.  Increase of freedom = 6 (7 - 1)  Improved (7 - 1)  Im	4 5 6 7 8 9	237 201811 237 201812 237 201901 237 201902 237 201903 237 201904 237 201905		132 131.254237 124 123.299438 117 116.338983 126 125.288136 119 118.327684 120 119.322034 129 128.271186	7 130 5 126 3 117 6 124 4 134 4 128 6 128	0.009556 0.021902 0.005682 0.010281 0.132448 0.072727 0.002114	1.28953 2.95570 2.0.76675 1.38745 3.17.87369 9.81442 4.0.28530	7 5 2 6 3 3			
STORE_WRR MonthD nCestomers controlCoat trialCoat percentage Diff Wake  7 207 201902 105 125 125 225193 124 0 010021 1 387456  8 227 201903 10 10 10 327204 1124 0 010021 1 387456  9 227 201904 101 10 327204 1124 0 010021 1 387456  9 227 201904 101 10 327204 1126 0 10021 1 387456  10 227 201904 101 10 327204 1126 0 10021 1 387456  10 227 201904 101 10 327204 1126 0 10021 1 387456  10 227 201904 101 10 327204 1126 0 10021 1 387456  10 227 201904 101 10 327204 1126 0 10021 1 387456  10 227 201904 101 101 101 201204 1126 0 10120777 88414478  We can observe that the t-value is much larger than the 95th percentile value of the 4-distribution for March and April - i.e. the in number of customers in the trial store in March and April - i.e. the in number of customers in the trial store, or in March and April - i.e. the in number of customers or the control store, trial stores, other stores and the 95th confidence level of control store is number of customers.  Visualizing the Sales of Trial and Control Store  Left oversite a more visual version of this by plotting the number of customers of the control store, trial stores, other stores and the 95th control store in other stores and the 95th control store in other stores of customers of the control store, trial stores, other stores and the 95th control store in the trial store and trial store in the store of control store in the trial store and trial store in the store of control store in the trial store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store in the store of control store in the store in store	Trial Period  Trial Store 237  **Discription:  **We can observe that the trial is store 88 is alignificantly different to its control store in the trial period as the trial store performance to the Store 237  **Discription:  **We can observe that the trial in store 88 is alignificantly different to its control store in the trial period as the trial store performance that store in the control store in the trial store in the control store in the trial store in the control store.  **Discription:**  **We can observe that the trial in store 88 is alignificantly different to its control store, that stores, other stores and the Store control store in the trial store in March and April is statistically greater than in the control store.  **Prevalence of customers in the trial and Control Store**  **enth create a more visual version of this by plutting the number of customers of the control store, that stores, other stores and the Stores create a more visual version of this by plutting the number of customers of the control store, that stores, other stores and the Stores created in the control store in the control store in the control store.  **Number of customers over the period**  **Trial Period**  **Discription:**  **Number of customers over the period**  **Trial Store 88 Control Store 237*  **Trial Period**  **Discription:**  **Discription:**  **Control store 233 for trial store 88 is alignificantly different to its control store in the trial period as the trial store performance in the store of the control store in two of the three trial months.  **Control store 233 for trial store 88 is alignificant difference in the sales and number of customers for two out of three trial months in the control store in the trial store store in the sales and number of customers for two out of three trial months in the sales and number of customers for two out of three trial months in the sales and number of customers in the period than pro-finil period however, all increased significant difference in the catagory manager, if some	significant Since there Hence, deg  degrees0 stats.t.; 1.894578	are 7 months in grees of freedom  fFreedom = 7  ppf(0.95,df=605061305	n the pre-tr	rial period ) Offreedom)				check w	hether the	e hypothesis is	statis
Visualizing the Sales of Trial and Control Store  Lefs create a more visual version of this by plotting the number of customers of the control store, trial stores, other stores and the 59% confidence level of control store's number of customers.  CoctrolStoreViz(measureOverTime, controlStore=237, trialStore=88, atdDev=atdDev, on="nCustomers over the period Trial Store: 88 Control Store: 237  Number of customers over the period Trial Store: 88 Control Store: 237  Number of customers over the period Trial Store: 88 Control Store: 237  Trial Period  Trial Period  Observation:  The results show that the trial in store 88 is significantly different to its control store in the trial period as the trial store performan outside the 5% and 95% confidence interval of the control store in two of the three trial months.  Conclusion  First we found control stores for trial stores as using pre-trial period data  Control store 233 for trial store 88  Control store 233 for trial store 88  Neacond the trial store 88 trial stores as using pre-trial period data  Control store 235 for trial store 88  Sales and number of customers increased for those months  However, for trial store 88 there was a significant difference in the sales and number of customers for two out of three trial months of the number of customers were increased for those months  However, for trial store 88 there was no significant difference in the sales but there was significant difference in the number of customers were increased for all three trial months  Naies was inceased from out of three trial months  Sales was inceased for one out of three trial months  Sales was inceased for one out of three trial months  This shows that, trial store 88 was able to altred significantly more customers in trial period then pre-trial period however, sale increased significantly. We need to check with the category manager, if some different startegy was applied at this store.	Visualizing the Sales of Trial and Control Store  of screate a more visual version of this by pioting the number of customers of the control store. It is stores, other stores and the 5% confidence level of control store's number of customers.  Number of customers over the period  Trial Store: 88 Control Store: 237  Number of customers over the period  Trial Store: 88 Control Store: 237  Trial Period  Tria	STORE_ 7 8 9	NBR MonthID  237 201902 237 201903 237 201904	nCustome	trolCust["Mcers controlCust 26 125.288136 19 118.327684	trialCust per 124 134	in([20190 rcentageDiff 0.010281 0.132448	tValue 1.387456 17.873693		4])]		
Trial Store: 88 Control Store: 237  Trial Store: 88 Control Store: 237  Trial Period	Trial Store. 88 Control Store: 237  Trial Period  Trial Pe	number of Visualizi Let's create 95% confid	customers in the ing the Sale a more visual vence level of con	he trial sto  s of Tria  version of to	al and Cont this by plotting t 's number of cus Time, contro	rol Store he number of stomers.	customers o	of the contr	ne control ol store, t	l store. rial stores,	other stores ar	nd the s
Observation:  The results show that the trial in store 88 is significantly different to its control store in the trial period as the trial store performance in the sales and number of customers for trial store 88. Control store in the assessed trial store 88. Control store in the sales and number of customers for trial store 88. Sales and number of customers increased for those months.  The we assessed trial store 88. Then we assessed trial store 88. Control store in the sales but there was significant difference in the sales but there was significant difference in the number of customers.  Number of customers were increased for all three trial months.	Others Trial Control Control State % confidence interval 201807 201808 201809 201810 201811 201812 201901 201902 201903 201904 201905 201810 201810 201811 201812 201901 201902 201903 201904 201905 201810 201807 201808 201809 201810 201811 201812 201901 201902 201903 201904 201905 201810 201807 201807 201808 201809 201809 201801 201801 201801 201901 201902 201903 201904 201905 201800 201807 201800 201800 201800 201801 201801 201801 201801 201801 201801 201901 201902 201903 201904 201905 201800 2	120							7	Period		
Conclusion  First we found control stores for trial stores using pre-trial period data  Control store 233 for trial store 77 Control store 155 for trial store 86 Control store 237 for trial store 88  Then we assessed trial store's performance in trial period  We found there was a significant difference in the sales and number of customers for two out of three trial months for trial store 88  Sales and number of customers increased for those months  However, for trial store 86 there was no significant difference in the sales but there was significant difference in the number of customers  Number of customers were increased for all three trial months Sales was inceasd for one out of three trial months  This shows that, trial store 86 was able to attract significantly more customers in trial period than pre-trial period however, sal increased significantly. We need to check with the category manager, if some different startegy was applied at this store.  Overall, impact of the new trial layouts at trial stores shows a significant increase in sales.	Conclusion  First we found control stores for trial stores using pre-trial period data  Control store 233 for trial store 77 Control store 155 for trial store 86 Control store 237 for trial store 88  Then we assessed trial store's performance in trial period  We found there was a significant difference in the sales and number of customers for two out of three trial months for trial store 88  Sales and number of customers increased for those months  However, for trial store 86 there was no significant difference in the sales but there was significant difference in the number of customers  Number of customers were increased for all three trial months Sales was inceased for one out of three trial months  This shows that, trial store 86 was able to attract significantly more customers in trial period than pre-trial period however, sal increased significantly. We need to check with the category manager, if some different startegy was applied at this store.	Observa The results	Control 95th % contro	nfidence inter	9 201810	Time	Period (yyyy	ymm) ol store in t	he trial pe	eriod as the		201 ormano
<ul> <li>We found there was a significant difference in the sales and number of customers for two out of three trial months for trial store 88</li> <li>Sales and number of customers increased for those months</li> <li>However, for trial store 86 there was no significant difference in the sales but there was significant difference in the number of customers</li> <li>Number of customers were increased for all three trial months</li> <li>Sales was inceased for one out of three trial months</li> <li>This shows that, trial store 86 was able to attract significantly more customers in trial period than pre-trial period however, sal increased significantly. We need to check with the category manager, if some different startegy was applied at this store.</li> <li>Overall, impact of the new trial layouts at trial stores shows a significant increase in sales.</li> </ul>	<ul> <li>We found there was a significant difference in the sales and number of customers for two out of three trial months for trial store 88</li> <li>Sales and number of customers increased for those months</li> <li>However, for trial store 86 there was no significant difference in the sales but there was significant difference in the number of customers</li> <li>Number of customers were increased for all three trial months</li> <li>Sales was inceased for one out of three trial months</li> <li>This shows that, trial store 86 was able to attract significantly more customers in trial period than pre-trial period however, sal increased significantly. We need to check with the category manager, if some different startegy was applied at this store.</li> </ul> Overall, impact of the new trial layouts at trial stores shows a significant increase in sales.	Concl First we fo Contro Contro Contro	und control store 233 for troll store 237 for	ores for tr rial store 7 rial store 8 rial store 8	rial stores using 7 6 8	control store in	two of the					
Overall, impact of the new trial layouts at trial stores shows a significant increase in sales.	Overall, impact of the new trial layouts at trial stores shows a significant increase in sales.	<ul> <li>We found</li> <li>88</li> <li>Sa</li> <li>However custom</li> <li>Number Sa</li> <li>This sh</li> </ul>	ales and number rer, for trial store ners umber of custom ales was inceaso	significant r of custom e 86 there v ners were i d for one o tore 86 wa	t difference in the ners increased for all increased for all out of three trial as able to attract	e sales and no for those mont ant difference in I three trial mo months t significantly r	hs n the sales b nths nore custom	out there w ners in trial	as signific	cant differe	ence in the num	ber of
		Overall, im	pact of the nev	w trial layo	outs at trial sto	eres shows a	significant i	increase i	n sales.			
												-