| Descrip This data set made, length These are the 1. hotel: Ty 2. lead_tim 3. arrival_d 4. arrival_d 5. arrival_d 6. arrival_d 6. arrival_d 7. stays_in_ 9. adults: N 10. children: 11. babies: N 12. meal: Ty 13. country: 14. market_s 15. distribut 16. is_repear 17. previous 18. previous 19. reserved 20. assigned room ty 21. booking 22. deposit_ 23. agent: IE 24. company designat 25. days_in_ 26. custome contract 27. adr: Ave 28. required 29. total_of_ 30. reservati 32. is_cance | earn.linear earn.tree is earn.metric earn. | the data the data king information the data set: If the reservation f days that elap frof arrival data Month of arrival month: Day of a ghts: Number of the segment of shildren abies coked. Categori rket segment of Booking distril alue indicating is: Number of production Code of room Code of room Code for the the company/entity Number of day of booking, ass of it; the as defined by estervation last is ate: Date at white ate: Date at wh | a set and a sion for a city hotel are lits, children, and/or on, City hotel or Reso psed between the end arrival date of weekend nights (Seek nights (Monday) are are represented in designation. In category bution channel. The lifthe booking name previous bookings the lumber of previous betype reserved. Code type of room assigned as the booking of that made the booking of that made the booking of special requests mustatus, assuming one of special requests mustatus was pooking was canceled the special requests mustatus. | Summary of indicate a resort hotel, and indicate and the number of the book of the book of the later of the later of the later of the later of the bookings of the later of the bookings of the bookings of the later of the booking. Some of all lodging transactions are perfectly the later of three categories: Contract - value of three categories of th | eans "Travel Agents" and el Agents" and "TO" meanguest (1) or not (0) he customer prior to the by the customer prior to of designation for anonymetimes the assigned room the moment the booking. This variable can paying the booking. ID is e it was confirmed to the when the booking has an ons by the total number of | e arrival date to stay at the hotel the hotel "TO" means "Tour Operans "Tour Operans "Tour Operators" current booking the current booking nity reasons. In type differs from the reasons assume three categories presented instead of customer allotment or other type of staying nights r) anceled by the customer |
|--|--|--|--|--|--|--|
| Resort Hotel Resort Hotel Resort Hotel Resort Hotel Resort Hotel | | ad_time arrival 342 737 7 | date_year arrival_da 2015 2015 2015 2015 | July July July July July July | 27 27 27 27 27 | te_day_of_month stays_in 1 1 1 |
| 1. Data cle 2. Bi-variak | (ploratio g to divide the aning: Detector | e data explorat t and treat miss Ve'll determine | e the interaction betw | | y building visualization to | |
| df.shape (119390, 3 Now let's see df.info() <class 'pa="" rangeindex<="" td=""><td>andas.core. : 119390 e</td><td></td><td>'rame'> 0 119389</td><td>lues:</td><td></td><td></td></class> | andas.core. : 119390 e | | 'rame'> 0 119389 | lues: | | |
| 2 lead 3 arriv 4 arriv 5 arriv 6 arriv 7 stays 8 stays 9 adult 10 child 11 babie 12 meal 13 count 14 marke 15 distr 16 is_re 17 previ 18 previ 19 reser 20 assic 21 booki 22 depos 23 agent 24 compa 25 days 26 custo 27 adr 28 requi 29 total 30 reser 31 reser dtypes: filling empty df = df.f | anceled _time val_date_ye val_date_mo val_date_da val_date_da s_in_weeker s_in_week_r s_in_week_r sdren es dren es try et_segment ribution_ch epeated_gue tous_cancel tous_bookin rved_room_t ng_changes sit_type ding_changes ding_changes sit_type ding_changes ding_chang | nannel est llations ngs_not_canc type type s g_list arking_space al_requests atus_date int64(16), MB ne column "cou | 119390 non | -null object -null int64 -null object -null object -null object -null object -null int64 -null object -null object -null object -null int64 -null object | n needed. | |
| #data df_countr df_countr df_countr #labels a lab=df_co colors = #plotting plt.figur plt.pie(o plt.axis() | shows the period highest ries pd. Date ries rename ries ["Perce and colors an | ercentage of w to lowest, mose aFrame (coun- e (columns={"centage"]=df_centage"]=df_centage"]=df_centage"]=df_centage"]=df_centage"]=dex[:35] palette('pase:(15, 10)) autopct='%1. | tries) | om according to their error Portugal (28%). of guests"}, inpl r of guests"]/sum | | |
| | | | 12.9% | | 28.0% | |
| Numb PRT GBR FRA ESP DEU | er of guests 21071 9676 8481 6391 6069 | Percentage 28.032621 12.872841 11.283027 8.502514 8.074129 | 8.1% | 3.4% 3.2% 2.5% ^{2.3%^{2.1}%} | | |
| DJI FRO SLE BFA GUY 166 rows × 2 | 1 1 1 1 columns | 0.001330 0.001330 0.001330 0.001330 | months? | | | |
| <pre>df = df.f df1=df[df order=["3" "F" "M" "3" "4" "5" "6" "1" df1["arri</pre> | fillna({"ch f["adults"] fanuary", february", farch", farch", fuly", fuly", fuly", fuly", foctober", foctober", foctober", focember", | #df["children": 0} #df["children": 0} |) en"]>0] ategorical(df1[" | | n"], categories=orde: | r, ordered= True) |
| A value is Try using See the caurning-a-v df1["arm # groupin df_resort df_city=d # countin df_resort | strying to .loc[row_i aveats in to view-versus rival_date_ ag by type ==df1.groupby ag visits p =1=pd.DataF col epd.DataFra | obe set on indexer, col_ the document s-a-copy month"]=pd. of hotel by("hotel") c("hotel").go per month Trame(df_reso tumns={"hotel tume(df_city.o | <pre>indexer] = value ation: https://p Categorical(df1[.get_group("Reso et_group("City H ort.groupby("arr l":"visits per m</pre> | re from a DataFrame instead andas.pydata.org/p "arrival_date_mont rt Hotel") otel") ival_date_month") onth resort"}) date_month")["hot | e. pandas-docs/stable/u th"], categories=ord ["hotel"].count()).re | er, ordered=True) |
| # cleaning months_20 months_20 months_20 months_20 # July and f_busy.ldf_busy.ldf_busy=0 # showing total=sum df_busy["df_busy["df_busy[" | col od.merge(df ag data, or 15=set(df_ 16=set(df_ 17=set(df_ 17=set(df_ 18 august a coc["July"] coc["August lf_busy/2 results a(df_busy[" visits per visits per | cumns={"hote." iresort1, d. inly 2016 is resort.group res | l":"visits per m f_city1, left_in complete, 2015 a pby("arrival_dat pby("arrival_dat pby("arrival_dat years and the ot c["July"]*(2/3) loc["August"]*(2 month city"])+su rt"]=df_busy["vi | onth city"}) dex=True, right_ir nd 2017 have the re_year").get_group e_year").get_group e_year").get_group her months only in /3) | ndex=True, how="outer lasts and the firsts 0(2015) ["arrival_date 0(2016) ["arrival_date 0(2017) ["arrival_date 0 (2017) ["arrival_date n 2, that's why I dis | months respectively e_month"]) e_month"]) e_month"]) |
| df_busy.p df_busy.r arrival_date_n | visits month anuary March April May June | p/m resort % 1.985051 2.810419 3.020612 3.269762 3.224462 2.757871 | e=(15,7.5)); s per month reso visits p/m city % 3.379388 4.484711 5.829219 6.766931 7.448245 7.143828 | | esort %","visits per | month city":"visits |
| Sept O Nov | June July August ember ctober ember | 2.757871 2.762099 2.955983 2.815855 3.219026 2.206116 2.396376 | 7.143828 4.874896 5.416082 6.697169 6.880181 3.928426 3.727293 | | | visits per mont visits per mont |
| 6 - 5 - 4 - 3 - 2 - | | | | | | |
| How lo The nights ar | re added on w | weekends and | | in_weekend_nights' | The stays of the stay o | |
| <pre><ipython-i <ipython-i="" a="" ca="" df_city<="" df_reson="" is="" pre="" see="" the="" try="" urning-a-v="" using="" value=""> The percenta</ipython-i></pre> | input-9-f33 s trying to .loc[row_i aveats in to riew-versus rt["nights input-9-f33 s trying to .loc[row_i aveats in to riew-versus rtying to .loc[row_i aveats in to riew-versus rtying to riew-versus rtying to reacts in t | Ba9070d920>: be set on indexer, col_ the document s-a-copy stayed"] = Ba9070d920>: be set on indexer, col_ the document s-a-copy ayed"] = df | 1: SettingWithCo a copy of a slice indexer] = value ation: https://p df_resort["stays 2: SettingWithCo a copy of a slice indexer] = value ation: https://p f_city["stays_in_ne hotels is obtained | pyWarning: te from a DataFrame instead andas.pydata.org/pin_weekend_nights pyWarning: te from a DataFrame instead andas.pydata.org/p weekend_nights"] depending on the nur | pandas-docs/stable/us"] + df_resort["stage. pandas-docs/stable/ustable | ser_guide/indexing. ys_in_week_nights"] ser_guide/indexing. week_nights"] |
| df_resort df_cityNS df_resort df_cityNS df_nights | NS=pd.Data =pd.DataFr NS=df_resc =df_cityNS _stayed=pd | Frame (df_restame (df_city orths/sum (df_cit; d.merge (df_restate)) | sort["nights sta ["nights stayed" _resortNS["night yNS["nights stay esortNS, df_city _index=True, rig | ed"])*100 NS, ht_index= True , how | 5()) | rhts staved v"·"nid |
| 25] | | | | | | nights stayed |
| 20 - | | | | | | nights stayed |
| The graph sh days How m # counting pplot1=df pplot2=df pplot3=df pplot3=df pplot3=df pplot3: ax1.pie (pax2.pie (pax3.pie (pax1.title ax2.title | any res ag cancella ["is_cance [df["hotel [df["hotel], ax2, ax3 pplot1, aut pplot2, aut pplot3, aut e.set_text(| ervation ations eled"].value "]=="Resort "]=="City He ded", "cancele sopet='%.0f%; copet='%.0f%; | tends to stay in the (S WERE CANC _counts() _Hotel"]["is_can otel"]["is_cance ed"] plots(1, 3, figs %',labels=labels %',labels=labels %',labels=labels %',labels=labels | celed? celed"].value_courts led"].value_counts ize=(20,10)));); | S () | → SS SS C4 C4 S5 |
| The graph shadays How m # counting pplot2=df pplot3=df pplot3=df pplot3=df pplot3-df grant ax1.pie (pax2.pie (pax3.pie (pax3.pie (pax3.title plt.show (pax | any res ag cancella ["is_cance [df["hotel for cancel | canceled has the language of the state of t | counts() Hotel"]["is_can otel"]["is_cance ed"] plots(1, 3, figs %',labels=labels %',labels=labels %',labels=labels el') ') not canceled rival_date_month al_date_month",y | celed? celed? celed!.value_courled"].value_counts ize=(20,10))););); Resort Hotel 72% 28% 28% | canceled canceled canceled canceled canceled canceled cancellation | nights stayed nights stayed nights stayed nights stayed st |
| The graph shadays How m # counting plot1=df pplot2=df pplot3=df #labels labels=["#plotting fig, (axlaxl.pie (pax2.pie (pax3.pie (pax2.pie (pax3.pie (pax2.title ax3.title plt.show (pax3.title plt.show (pax3.pie (pax3.pie (pax3.title plt.show (pax3.pie (pax3.pie (pax3.title plt.show (pax3.title plt.show (pax3.title plt.show (pax3.pie (pax3.title plt.show (pax3.title | any res any res ag cancella ["is_cance [df ["hotel not cancel ", ax2, ax3 plot1, aut plot2, aut plot3, aut eset_text(eset_tex | ervation ations aled"].value "]=="Resort "]=="City Ho add", "cancel appet='%.0f% appet='%.0f | counts() Hotel"]["is_can otel"]["is_cance ed"] plots(1, 3, figs %',labels=labels %',labels=labels %',labels=labels el') ') not canceled rival_date_month al_date_month",y | celed? celed? celed!.value_courled"].value_counts ize=(20,10))););); Resort Hotel 72% 28% 28% | canceled canceled canceled canceled canceled canceled canceled canceled canceled and not canceled | nights stayed nights stayed nights stayed nights stayed st |
| The graph shadays How m # counting polotical of polotic | any res any res any res ag cancella any ax2, ax3 aplot1, authority, author | canceled has the land that illigrarival in add [0] / ddd ["] / canceled has the land that illigrarival in add [0] / ddd ["] / cancellation in add [0] / | April May counts () Hotel"] ["is_can otel"] ["is_cance ed"] plots (1, 3, figs %',labels=labels %',labels=labels %',labels=labels el') not canceled rival_date_month al_date_month al_date_month al_date_month yrrival_date_mont rival_date_mont entry al_date_month al_ | June July arrival_date_end % ", y="canceled % ", hue h", "canceled % "]]. June and the percentage ", y="canceled % ", hue h", "canceled % "]]. | August September orital and May. gineering statch for the correlation be a cancelled and not concern and a cancelled and so that influence a cancella arach for the correlation be a cancel arach for the cancel area. | nights stayed nights stayed nights stayed nights stayed stayed nights stayed stayed nights stayed nights stayed stayed nights ni |
| The graph shodays How m # counting political and politic | any res any res any res ag cancella ag c | canceled has the land that it [["arrival ded [0] / ddd ["] cancellation and [0] / ddd ["] cancellation and [0] / ddd ["] canceled has the land that it [["arrival ded [0] / ddd ["] cancellation and [0] / ddd ["] canc | April May counts () Box Were Cance S Were Cance S Were Cance Counts () Local [] ["is_cance dotal ["is_cance dota | City Hotel for 3 or 4 day celed? celed? celed!.value_counts ize=(20,10))););); Resort Hotel 72% ages"].loc[month]. and the percentage ",y="canceled %",hueah","canceled %",lueah","canceled %"]]. thone are the variables are the variables are the months, or fill displayed by April 1. | August September anceled and not concerning sum() e canceled and not concerning sort_values(by="canceled"); sort_values(by="canceled"); ing the empty spaces. | nights stayed in index stayed |
| The graph shadays How m # counting polotical | any res any res any res ag cancella ["is_cance [df["hotel not cancel ax2, ax3 plot1, aut plot2, aut plot3, aut caset_text(caset_te | canceled has the land that it [["arrival ded [0] / ddd ["] cancellation and [0] / ddd ["] cancellation and [0] / ddd ["] canceled has the land that it [["arrival ded [0] / ddd ["] cancellation and [0] / ddd ["] canc | April May counts () Box Were Cance S Were Cance S Were Cance Counts () Local [] ["is_cance dotal ["is_cance dota | City Hotel for 3 or 4 day celed? celed? celed!.value_counts ize=(20,10))););); Resort Hotel 72% ages"].loc[month]. and the percentage ",y="canceled %",hueah","canceled %",lueah","canceled %"]]. thone are the variables are the variables are the months, or fill displayed by April 1. | August September orital and May. gineering statch for the correlation be a cancelled and not concern and a cancelled and so that influence a cancella arach for the correlation be a cancel arach for the cancel area. | nights stayed in index stayed |
| The graph shadays How m # counting polotical | any res any res any res ag cancella ag inscalla ag i | ervation ations ations ations aled"].value aled"].value aled", "cancel aled", "cancel aled", "cancel aled", "cancel aled", "cancel aled at l atil [["arrival add [0] add (0] add (0) add | tends to stay in the G S Were CanC _counts() _Hotel"]["is_cance dotal"]["is_cance ed"] plots(1, 3, figs %',labels_labels %',labels=labels %',labels=labels %',labels=labels %',labels=labels el') ') counts the valu _date_month", yis th]=ddd["percent percentages"]*10 ons is obtained rival_date_month al_date_month yrrival_date_mont counts the valu _date_month", y rrival_date_mont counts the valu _date_month", y rrival_date_mont date_month", y rrival_date_mont counts the valu _date_month", y rrival_date_mont atrix in order to whice yould be when we order counts the valu _date_month", y rrival_date_mont counts the valu _date_month", y rrival_date_month al_date_month object"]["index" counts the valu _date_month object counts the valu _date_month obje | June July June July June July Resort Hotel Ages"].loc [month]. and the percentage ", y="canceled %", hue- h", y="canceled %", hue- h", "canceled %", hue- h | August September canceled and not consume is canceled"; sum() canceled and not consume is canceled"; sort_values (by="canceled"); sort_values (by="canceled") arch for the correlation being the empty spaces. | City Hotel City Hotel not canceled acceled \$", as cendings anceled \$per month is celed \$", as cendings tion. For that we need to etween the variables. |
| The graph shadays How m # counting political filt political filt political filt political filt political for month with the columns of the | any results any results any results are an ella grancella gra | ## Company | tends to stay in the of tends to stay in the of tends to stay in the of tends of the other was a stay in the of tends of the other was a stay in the of tends of the other was a stay in the other was | City Hotel for 3 or 4 da celed? celed? celed! value_counts ize=(20,10)) ; ; celed! value_counts ize=(20,10)) ; ; Resort Hotel re | August September oril and May. Gineering Stations? Ser canceled and not conceed and set of a concellation accounts ()).reset_ind sum () S | October November Control of the November Control of th |
| The graph shadays How m # counting plotlading plotlading plotlading plotlading plotlading plotlading plotlading plotlading shadays plotlading sh | any res any res any res any res any res aga and ella | ## Complementally | April May Counts (1) Flote (1) (1) (1) (2) (2) (2) (2) (2) (3) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4 | City Hotel for 3 or 4 da celed? celed. cele | August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. August September arrival date of the correlation being the empty spaces. | The property of the property o |
| The graph shodays How m # counting pplotled and should be should | any resolutions and the that are string taken for date that ar | ## Complet mostly: ## Cartion ## Cartions | April May Counts () Footel "] ("is can counts () Footel "] ("is can counts () Footel "] ("is can counts (), 3, figs \$ ', labels = labels \$ ', lab | | ### August September ### Canceled ### August September ### Canceled #### Canceled ##### Canceled ##### Canceled ##### Canceled ##### Canceled ###### Canceled ################################## | and and a second a |
| The graph she days How m # counting political file plots and file plots and file and file and file file file file file file file file | any res any | decople mostly: ervation trions tri | April May Source Cance Source Cance Counts () Botel | June July Present Hotel Table To G count and a count | ### August September ### Canceled ### August September ### Canceled #### Canceled ##### Canceled ##### Canceled ##### Canceled ##### Canceled ###### Canceled ################################## | City Hotel October November C Since lod per month is sayed and the say of t |
| The graph shadays How m # counting plottled from the graph shadays How m # counting plottled from the graph shadays # and the graph shadays | any res any | decople mostly ervation ations tions tio | April May cancellations in total cleaning are strick in order to which value, doing that, py und be seen and the reace and canceled counts the value and canceled and canceled and canceled counts the value cleaning are strick in order to which value, doing that, py und be when we orde reset_index () i) and canceled rival_date_month al_date_month al_date_month al_date_month al_date_month al_date_month al_date_month al_date_month al_date_month al_date_strice counts the value cleaning are strick in order to which value, doing that, py und be when we orde rival_date_month al_date_month al_date_month al_date_month al_date_month cleaning are strick in order to which value, doing that, py und the variable that und the var | The Hotel for 3 or 4 de cele ed? Cele ed. Cel | August September August September Canceled and not or Counts () reaceled and not or Count | City Hotel October November C Since lod per month is sayed and the say of t |
| The month was a date of the columns | any res any | ervation ervati | April May Counted () Both () | June July June July June July June July June July June July A Resert Hotel 7256 Resert Hotel 7276 Resert Hotel 7277 Resert Hotel 7277 Resert Hotel 7277 Resert Hotel 7277 Resert Hotel 727 | August September accessed Selfations? Ser. Gardel St. 100 South 18 (1) - reset 1 | City Hotel The concept of the conce |
| The month ware and a second trime and a second trime and a second trime and a second trime are going to the columns of the col | any resident and any | ata exple alternation ervation ervation ervation cross and a servation attention | tends to stay in the company of the | The product of the second of t | and the control of th | City Hotel The concept of the conce |
| The graph sh days How m for countri population for countri population for countri population for countri for countri population for countri for coun | any res any res any res any res any res any cancel any cance | ata explusions of the strumber of concelled and and and and and and and and and an | and the state of the cancel and the state of the state of the cancel and the state of the cancel and the state of | June July Selection of Canace Control of Canace | August September of the control of t | City Hotel City Hotel City Hotel MX AN |
| The graph shadays Which is a construction of the construction of | any res any res any res any res any res any res any carea any res any carea any | ata explanation and a content of the | April May ancellations in total cleaning ar strict in order to whice and to be able | About the service of | August September August September August September Size Connect Colffy Size Connect Colff | City Hotel The control of the contr |
| Which is a contribution of the contribution of | any res any | ### A CONTRICT OF THE PROPERTY | April May and and and accuracy April May and and accuracy and and accuracy accuracy and accuracy | June 1949 Selected? Collected? Collecte | August September August September August September Size Connect Colffy Size Connect Colff | City Hatel City Hatel The cancel of the ca |

