**E-commerce Website**

Tan Eu Liong (Jerome)

Funnel Analysis



**Problem Statement**

You are looking at data from an e-commerce website. The site is very simple and has just 4 pages: The first page is the home page. When you come to the site for the first time, you can only land on the home page as a first page. From the home page, the user can perform a search and land on the search page. From the search page, if the user clicks on a product, she will get to the payment page, where she is asked to provide payment information in order to buy that product. If she does decide to buy, she ends up on the confirmation page.

The company CEO isn't very happy with the sales and, especially, sales coming from new users. Therefore, she asked you to investigate whether there is something wrong in the conversion funnel or, in general, if you could suggest how conversion rate can be improved.

**Purpose**

A full picture of funnel conversion rate for both desktop and mobile. Unlock crucial insights on what the product team should focus on in order to improve conversion rate as well as anything you might discover that could help improve.

**End Result**

Provide quick conversion with graphs on insight in order to support data driven marketing.

**Success criteria**

* Digestible visualisations uncovering funnel conversion insights with latest data available
* Able to make better decisions & spend on key areas for improvement of conversion.
* Eliminate assumptive or bias reasoning of low conversion rate, to save 20% of business time and reinvest it to value added activity.

**Resources**

In place of the domain knowledge required, I have used an article showing online consumer statistics to guide my thought process with coming up with suggestions to improve conversion rate

* ***Invespcro***

*Online Consumer Shopping Habits and Behavior*

<https://www.invespcro.com/blog/online-consumer-shopping-habits-behavior/>

* ***Smartinsights***

Ecommerce Funnel conversion rates

<https://www.smartinsights.com/ecommerce/ecommerce-analytics/ecommerce-funnel-conversion-rates/>

**Environment Used**

**Azure Data Studio**

A cross-platform database tool for data professionals using on-premises and cloud data platforms on Windows, macOS, and Linux. Azure Data Studio offers a modern editor experience with IntelliSense, code snippets, source control integration, and an integrated terminal.

**Importing Libraries & Datasets**

* **Libraries**

|  |  |
| --- | --- |
| **numpy** | mathematical functions |
| **pandas** | data manipulation |
| **plotly** | funnel graphs |
| **calendar** | date conversion |

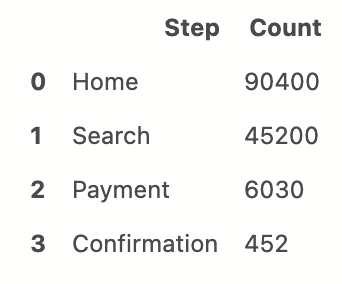
* **Datasets**

|  |  |
| --- | --- |
| **home\_page\_table.csv** | user\_id, home\_page |
| **search\_page\_table.csv** | user\_id, search\_page |
| **payment\_page\_table.csv** | user\_id, payment\_page |
| **payment\_confirmation\_table.csv** | user\_id, payment\_confirmation\_page |
| **user\_table.csv** | user\_id, date, device, sex |

These datasets have no null values and no outliers. Hence we shall proceed into data exploration.

**Count of Visitors on Each Page**

Our first milestone is would be to get an overview of how the visit rates look like as customers go deeper into the sales funnel. We will be able to immediately draw some insights on the conversion of our site. Counting the number of records for each dataset tells us the number of visits to each page.



page\_count = pd.DataFrame([['Home', home['user\_id'].count()],

['Search', search['user\_id'].count()],

['Payment', payment['user\_id'].count()],

['Confirmation', confirmation['user\_id'].count()]],

columns = ['Step', 'Count'])

page\_count

* Similar to the home page, **user\_table** also contains 90400 observations.
* with info() we verified that there are no null values in each of the datasets

**Visualising Count of Visitors by Funnel Analysis**

With the above results, we plot a funnel graph to get a visual representation of how the website conversion looks like.

fig = go.Figure(go.Funnel(

y = ['Home', 'Search', 'Payment', 'Confirmation'],

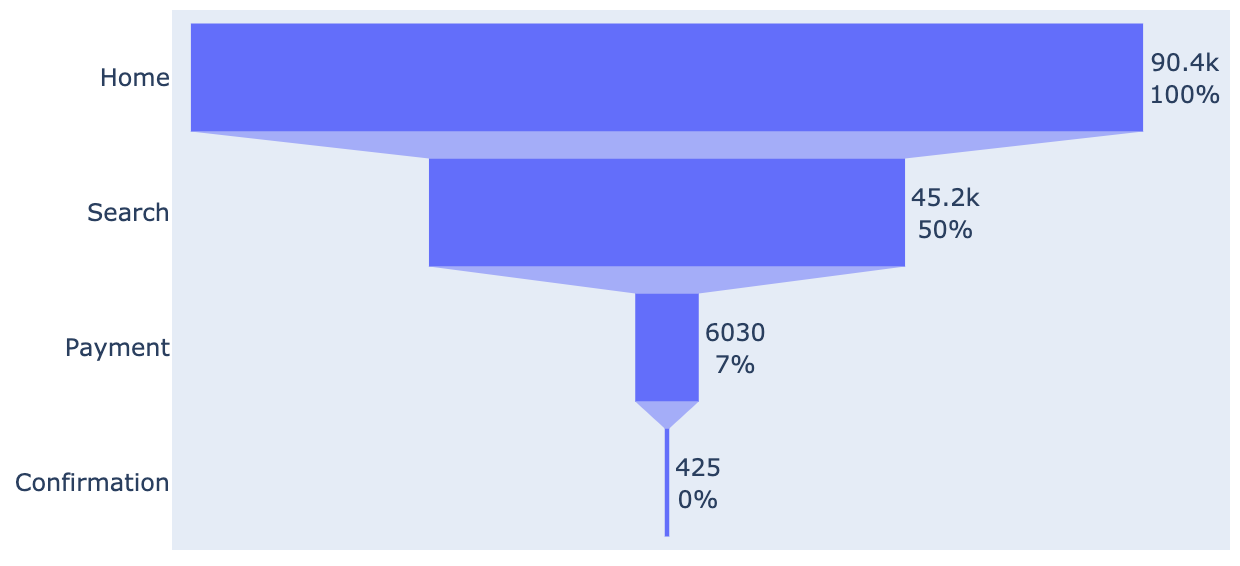
x = [90400, 45200, 6030, 425],

textposition = 'outside',

textinfo = 'value + percent initial'))

fig.show()

*Funnel chart for no. of visits per page*



The funnel chart helps visualise a linear process of the e-commerce webpage that has 4 sequentially connected stages:

1. **Home: 100%**

This represents the total sessions recorded in our website

1. **Search: 50%**

Half of the total sessions recorded went on to search for a product

1. **Payment: 6.7% (rounded off to 7)**

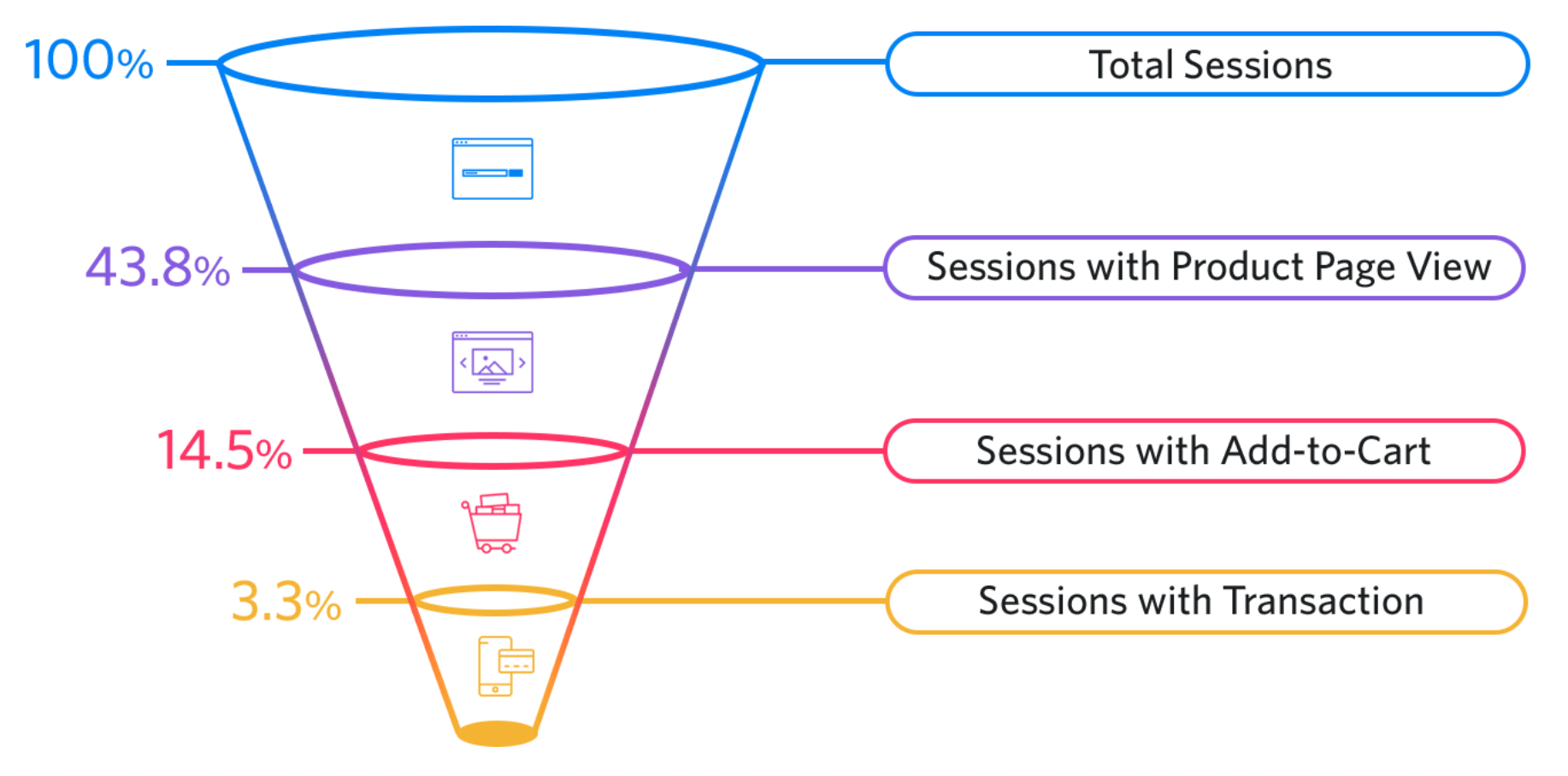
Only 6.7% of the total sessions ended up at the payment page

1. **Confirmation: 0.5% (rounded off to 0)**

Only 0.5% of the total sessions purchased a product

**Comparing our funnel with the average conversion rate**

*Average funnel chart from SmartInsights.com*



By comparing the conversion percentages between our website and the average website, our funnel is performing worse in the purchase and confirmation pages.

**Analysis**

* Our search page got 50% of total visits, better than the average visit rate of 43.8%
* Our purchase page only got 7% of total visits, below the average visit rate of 14%.
* Our confirmation page only got 0.5% of total visits, below the average rate of 3.3%

**Possible reasons for trend in higher searches yet low conversion rate:**

**High search rate**

* Home page better designed than competitors
* Shows better promotions

**Low conversion**

* Product(s) searched not to their preference
* Product(s) found but were pricey compared to competitors
* Uncertainties about the product/purchase related concerns
* Issues with shipping costs
* Issues with mode of payment

**Suggestions**

* Offer more competitive prices
* Have at least one Customer Service Rep on standby to answer to customer’s concerns
* Offer a wider variety of brands and products
* Offer cheaper or free shipping (80% of surveyed customers are more likely to purchase a product when offered free shipping)
* Incentivise by providing coupons for purchase
* Ease in return and exchange policy
* Offer more payment options

We now have a general insight on the conversion rate of our website. With some data manipulation, we are able to draw further observations from new variables.

Here we will shorten the column names in each dataset for better readability(e.g., ’home\_page’ to ‘Home’’). The 5 tables are joined using outer joins as 1 dataset named *flow* shown in the figure below

# shorten names of columns

df = df.rename(columns = {‘home\_page' : ‘Home'})

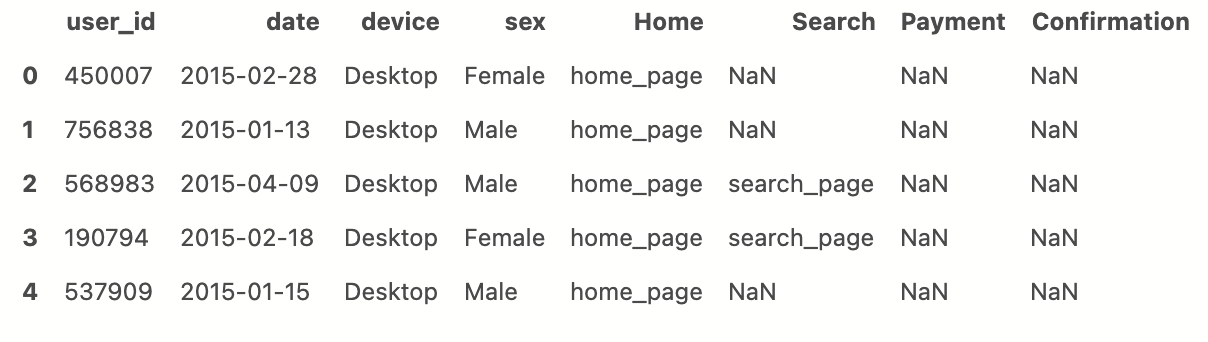
# merge all the tables

flow = df.merge(df1, how = 'outer', on = 'user\_id').merge(search,

how = 'outer', on = ‘df2’)

flow.head()

*flow: Result of merged table using outer joins*

****

* We used to *nunique()* check if there are any unique values in user\_id and found that all the values in this column are unique. Therefore all 90400 records shall be treated as unique users.

Using *flow* dataset, we produced another similar dataset *flow2*. In this new dataset, we created dummy variables for all the categorical columns, thus the values are converted to numerical values for plotting the correlation table.

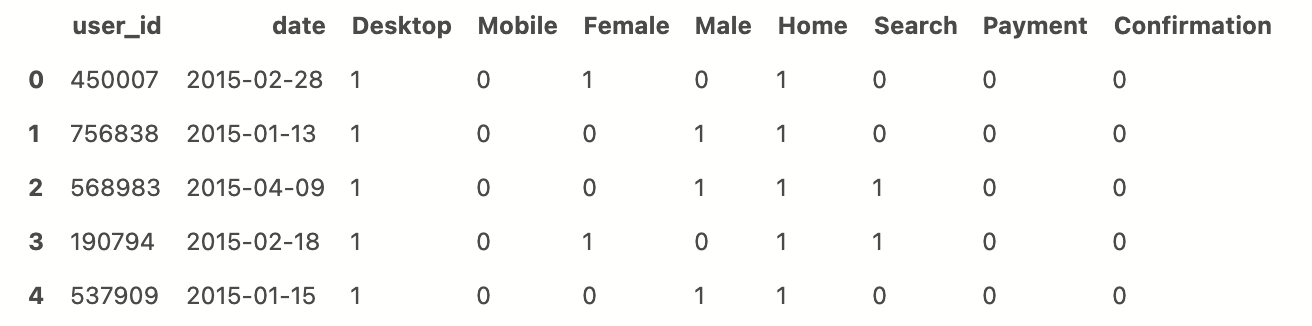
# dummy variables on all features : flow2

flow2 = pd.get\_dummies(data=flow,

columns = ['device', 'sex', 'Home', 'Search', 'Payment', 'Confirmation'],

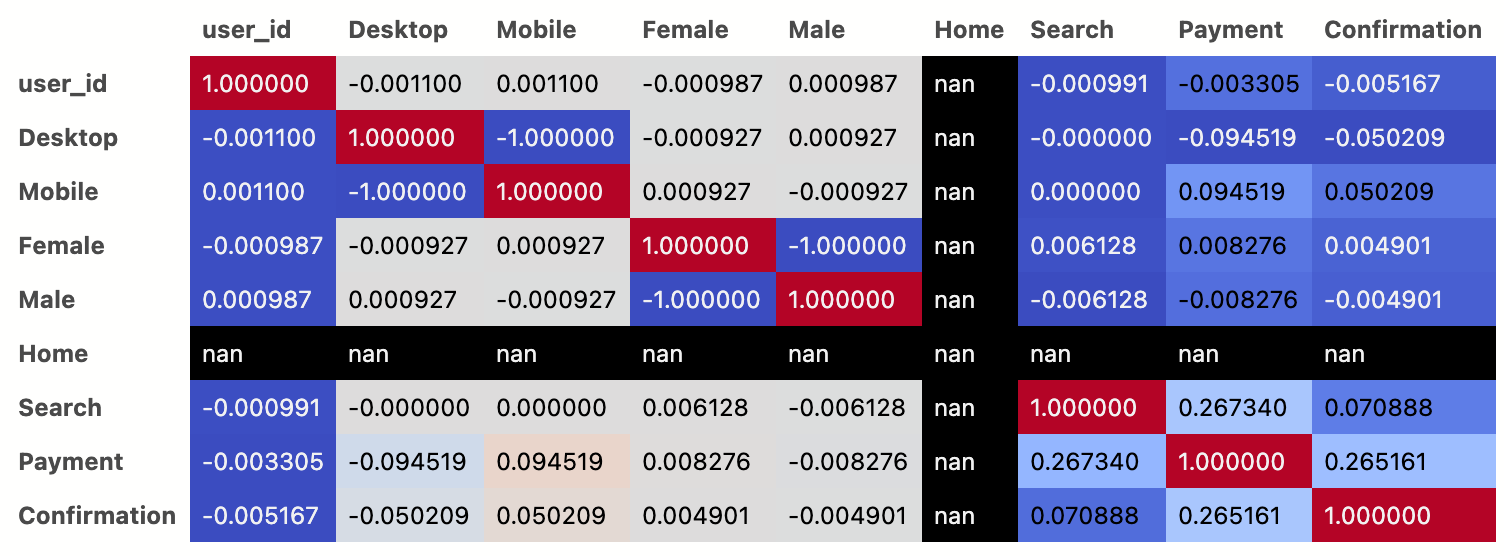
dummy\_na = False)

*flow2: Similar table with dummy variables (columns have been renamed for better readability)*



**Overview of Correlation**

*flow2 correlation table*

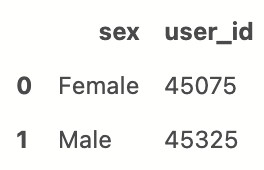


Red: positively correlated

Blue: Negatively correlated

* There is little correlation

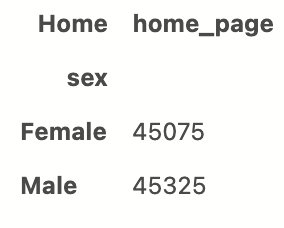
**Comparing Visits by Gender**



* The distribution of male and female visitors is almost equal

**Comparing Visits per Page by Gender**

For ease of repeatability we create a functions like this that plots pivot tables through out analysis

Result: 

# This function uses sex as y-axis and takes in input of page for x-axis columns

def gender\_pivot(page):

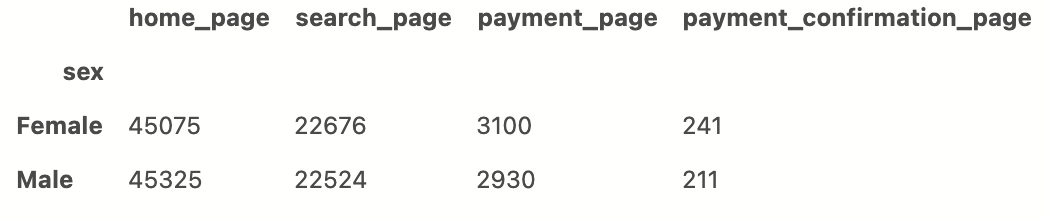
gender\_page = flow.groupby(['sex', page]).user\_id.count().reset\_index()

gender\_pivot = gender\_page.pivot(index='sex', columns = page, values = 'user\_id')

return gender\_pivot

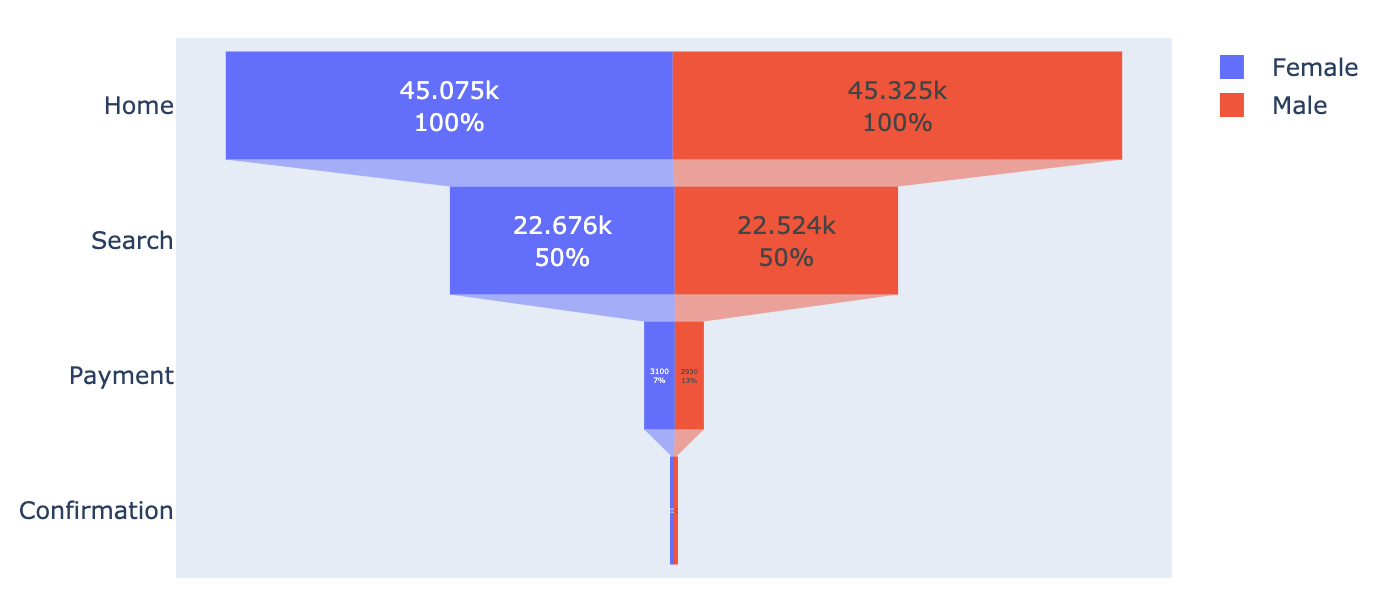
gender\_pivot('Home')

The same function is used for search page, payment page & confirmation page. All resulting tables are joined to produce this table. We plot this table as a funnel chart using graph\_objects from plotly.

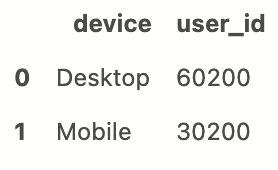


* The amount of visitors base on gender, is almost similar across each page

*Representation by funnel chart*

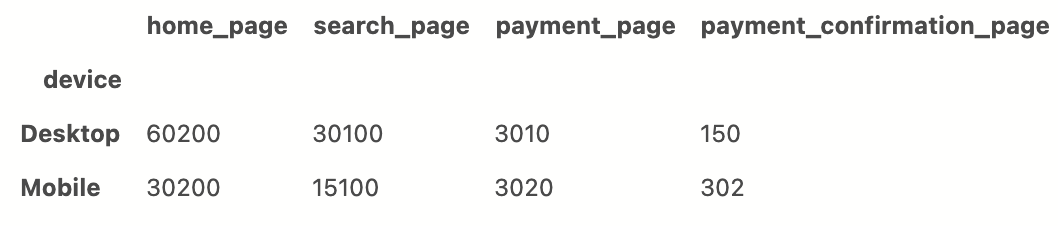


**Comparing Visits by Device**

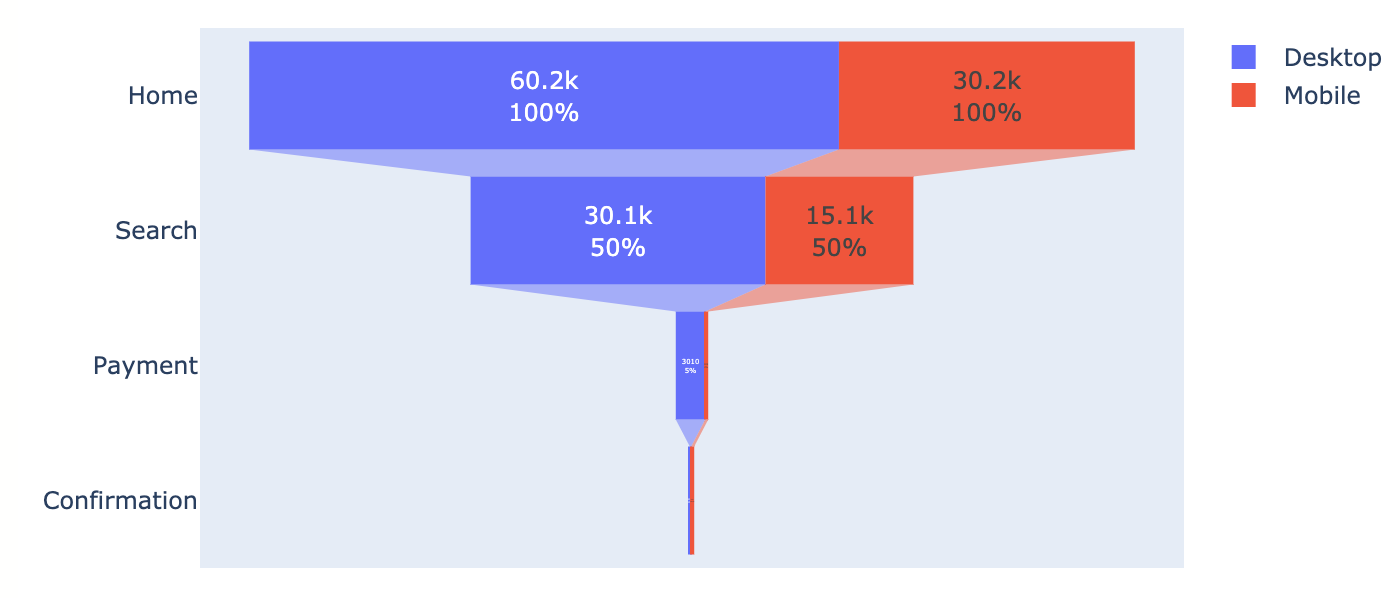


* Desktop has almost twice as many visits as Mobile

**Comparing Visits per Page by Device**

****

*Representation by funnel chart*

****

**Analysis**

Even though there are around twice as much traffic with desktop, the amount of visits to the payment and confirmation page is more for mobile than desktop.

This pattern supports the statistics offered by the aforementioned article from [invespcro.com](http://invespcro.com) which states:

* 1. *The time spent on retail sites by device are:*
  + *PC: 49%*
  + *Smartphone: 37%*
  + *Tablet: 14%*
  1. *Devices used to make purchases:*
  + *Smartphone: 50%*
  + *Tablets: nearly 60%*

**Some possible reasons for this trend:**

Mobile:

* Convenient payment with more modes of transaction using mobile device   
  (payment could be done via linking with other apps like mobile banking applications)

Desktop:

* Visitors utilise multiple tabs when searching for products (Switching between tabs is more conveniently done using desktop devices.)
* Often offer better viewing of products than mobile devices which have to be scaled to fit a smaller screen. (More scrolling involved in mobile devices)

**Suggestions**

**Improving Sales Across Each Device Type**

We have observed that even though more searches are done via desktop, purchases are usually done using mobile devices. However we should still improve the conversion rate of desktop users because those who left the website might never return to purchase using their mobile devices.

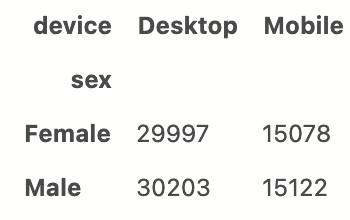
Mobile:

* Further improve sales by directing more marketing efforts towards mobile users on social media sites. (Statistics have shown that 84% of online shoppers refer to at least one social media site for recommendations before shopping online)
* Improve the mobile web design across all pages for better user experience when visitors are searching for products

Desktop:

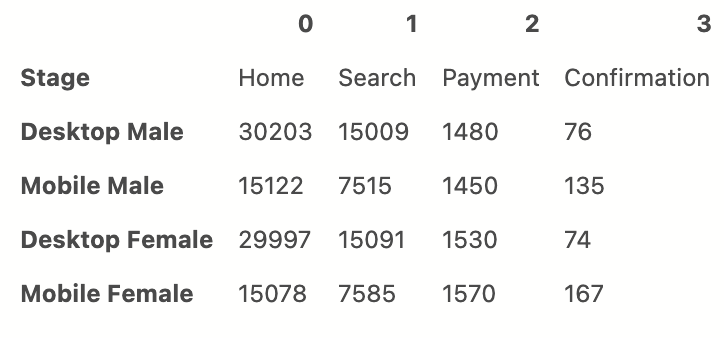
* Improve the web interface for easier transactions with more modes of payment so that visitors will not need to switch between devices when making a purchase

**Comparing Visits Base on Gender & Device**

****

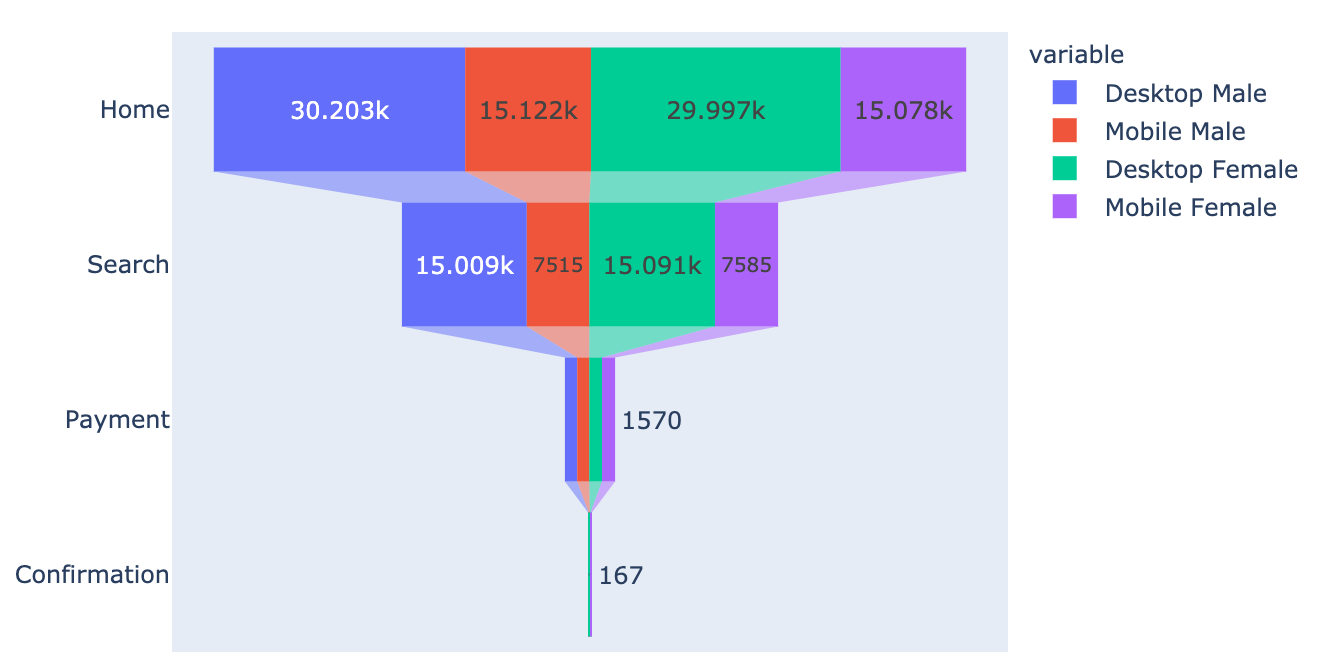
* Male and Female visitors contribute to similar amount of visits on desktop
* Male and Female visitors contribute to similar amount of visits on mobile

**Comparing Visits per Per Page base on Gender & Device**



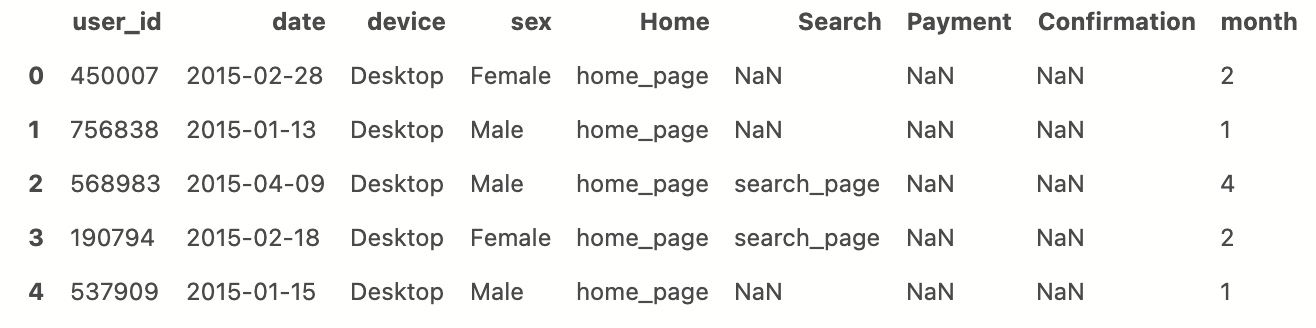
* Male and Female visitors contribute to similar amount of visits on desktop across all pages
* Male and Female visitors contribute to similar amount of visits on mobile across all pages

*Representation by funnel chart*

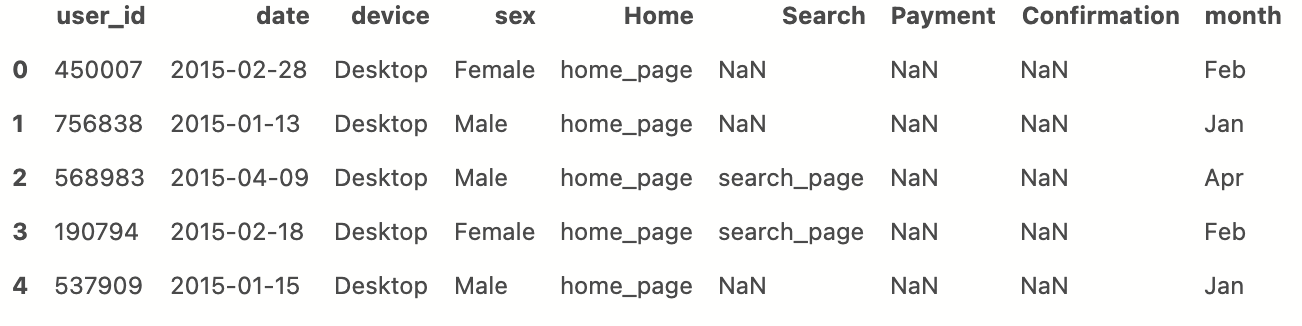


**Customer Churn**

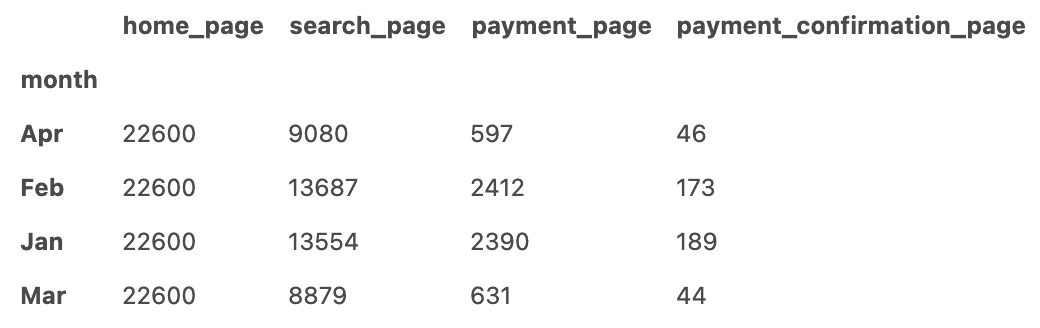
Customer churn, is the rate at which customers stop doing business with an entity. For our churn analysis, a new column ‘month’ is created using DatetimeIndex() . This will be used to measure our churn rate with.

****

The months are represented by numbers, instead of using this we convert them to their respective names.

****

**Churn Rate Across All Pages**

****

* January: most sales at 189
* February: slightly less sales at 173
* March: sharp drop is sales at 44
* April: remains low at 46

**Suggestions**

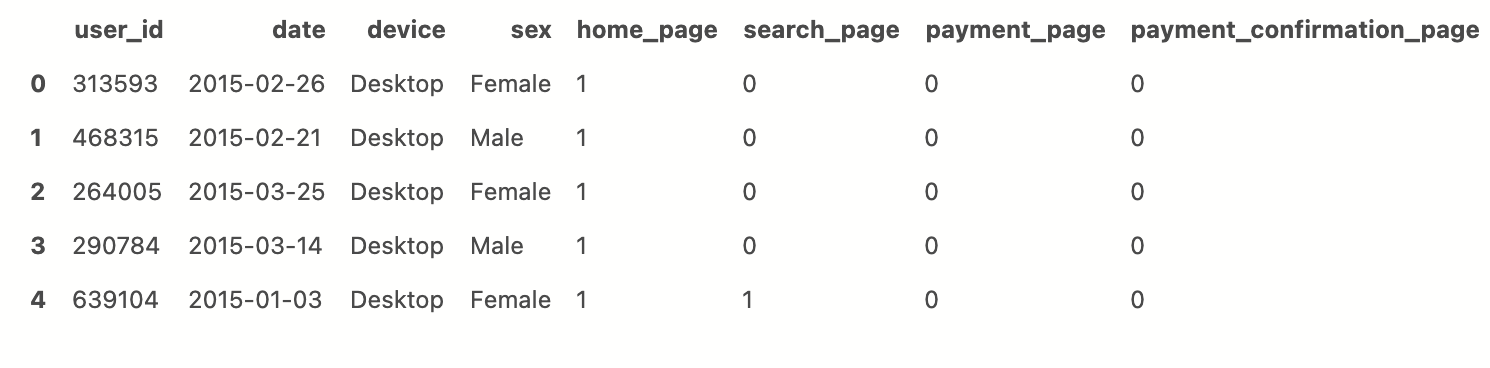
**Improving Sales Across Months**

There is a higher volume of visits to the home page on months with less sales like March and April. It seems that more awareness has been generated during this period. However most of the visitors leave before the purchase page. We could increase the chances of them making a purchase by offering coupons when profiting the website.

**Additional Activity With PowerBI**

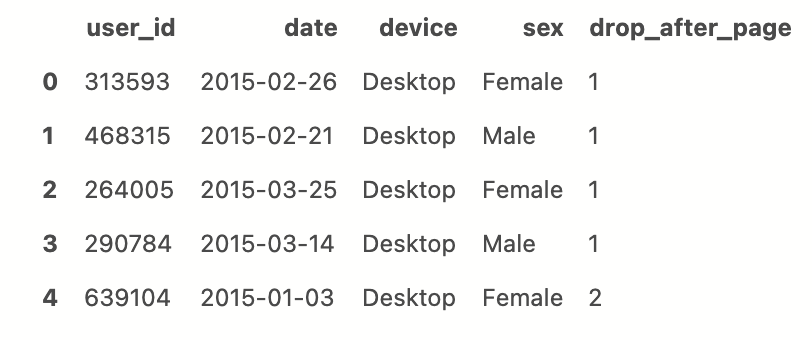
I have also experimented a different way in handling this problem statement. In this attempt, I

merge the 5 tables and used one-hot-encoding for the 4 stages of our funnel.



This table shows us the records of visitors who have reached each page.

We are able to combine these numerical columns into one to get the total pages visited before dropping out of the funnel. As we can see, the record of user\_id ‘639104’ visited the home and search page, therefore it it represented by ‘2’



1 = only visited home page

2 = visited home page & search page

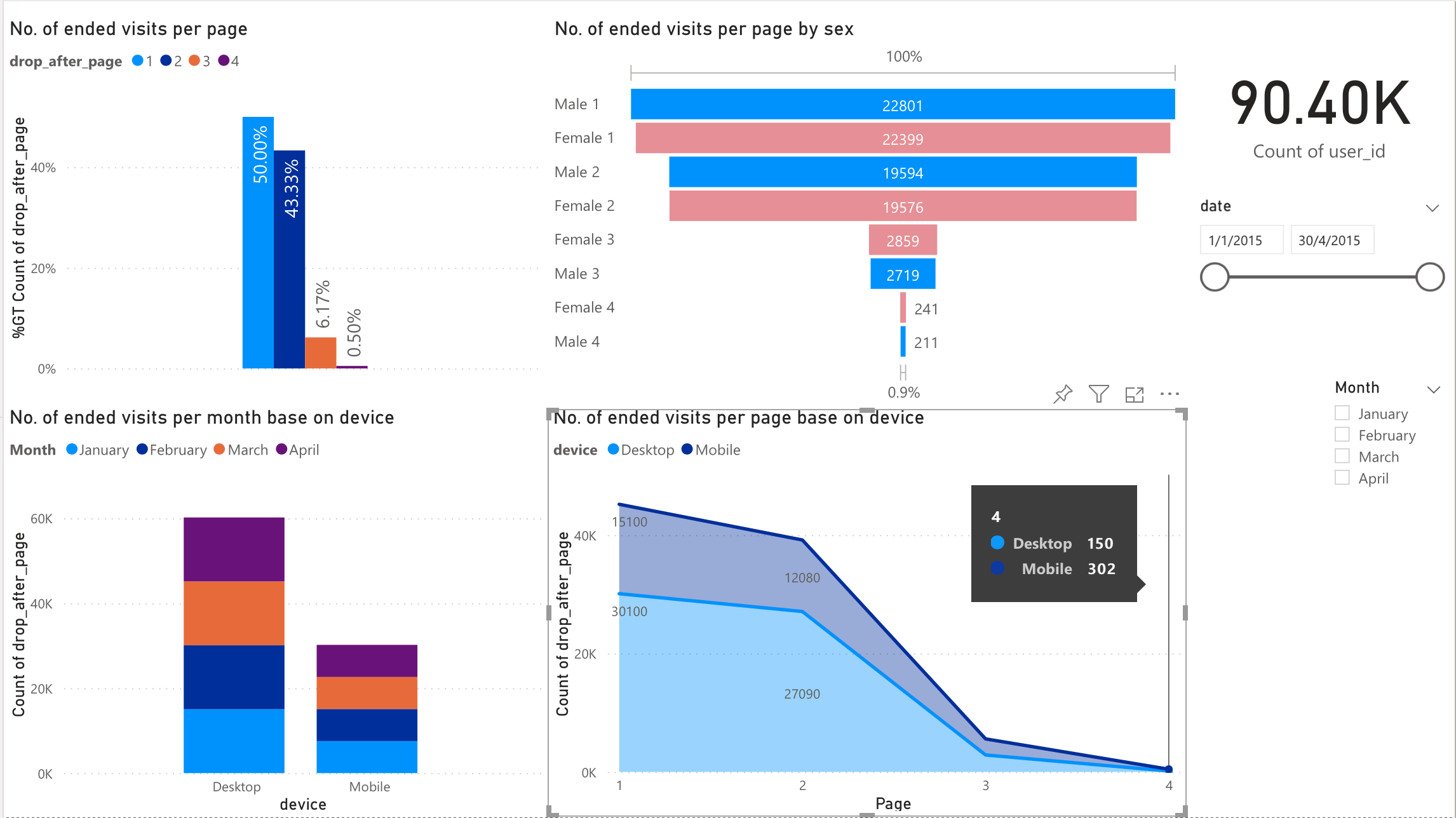
3 = visited home page, search page & payment page

4 = visited home page, search page, payment page & payment confirmation page

The benefit of handling this problem statement this way is that it makes the table simple and clean with only 5 columns.

**Some visualisations made using this method on PowerBI**

<https://app.powerbi.com/view?r=eyJrIjoiNDcwZTVlYjQtNzJjNS00M2UxLTkyNjUtZGE3ZjY0YTcxYzA2IiwidCI6IjMxMjUxOTA2LTE1OTItNDAwNy1iMTQ3LTE5Yzk2M2U2NzliOSIsImMiOjEwfQ%3D%3D>

****