Introduction and Traffic Analysis

Part I

Clickstream Analysis

A clickstream is a stream of events that represent **user actions** on a website or a mobile applications.

Today, **websites and mobile applications** have become the digital storefronts of every eCommerce company. Clickstream data is critical for business process insights like customer traffic analysis, marketing campaign management, market segmentation, sales funnel analysis, and so on.

Clickstream data is critical for business process insights like customer traffic analysis, marketing campaign management, market segmentation, sales funnel analysis, and so on.

In marketing campaigns, there can be different **optimization objectives**: for example, increasing sales monetization, improving customer retention, or extending brand reach. Insights need to be extracted from raw data consisting of web tracking events.

Automation

The objective of automating the clickstream analysis is to enable the marketing team to get answers for **more than 80%** of their questions.

They can approach the technical team **only** when they need details for any specific query.

This kind of automation reduces the time required to take decisions and management people are well informed by the application. They can use the application any time anywhere.

Automation Approach

Once the data is stored in any place by any technology, **R program** acquire, process and aggregate the data. The output of the R will be stored as **data cube** and that will be given as input to the **Power BI**.

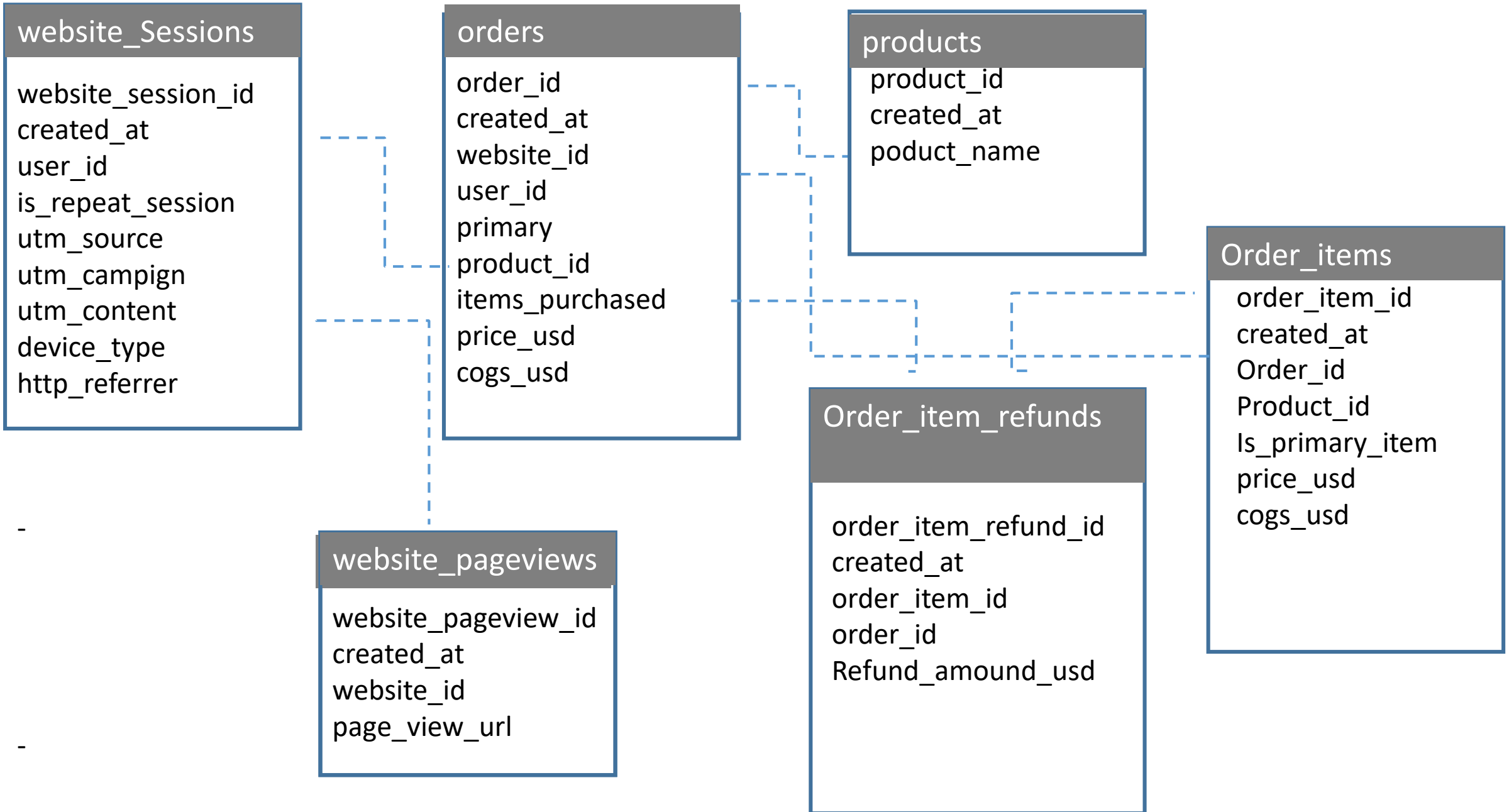
Now power BI creates reports and dashboards, with the input data, that the marketing team can use for their operations.

How the automation is to be done is, R program can be scheduled to run by the **operating system** and that will produce the required result. Power BI can also be scheduled to **refresh the data** for its reports and dashboards.

Data

The data, I have obtained, was created by **John Pauler** of Maven Analytics and I have got it through the udemy course **“Advanced SQL: MySQL Data Analytics & Business Intelligence”** by Maven Analytics, John Pauler.

Schema


website_Sessions
website_session_id
created_at
user_id
is_repeat_session
utm_source
utm_campign
utm_content
device_type
http_referrer
orders
order_id
created_at
website_id
user_id
primary
product_id
items_purchased
price_usd
cogs_usd
products
product_id
created_at
poduct_name
Order_items
order_item_id
created_at
Order_id
Product_id
Is_primary_item
price_usd
cogs_usd
Order_item_refunds
order_item_refund_id
created_at
order_item_id
order_id
Refund_amound_usd
website_pageviews
website_pageview_id
created_at
website_id
page_view_url

-

-

Traffic Analysis

The easiest way to utilize clickstream data is to see where a website is getting traffic from.

This analysis provide insights into who visits our websites, where they come from and what content they are interested.

The website sessions coming from different sources are analysed and their breakdowns are shown.

This matrix shows the breakdown of the sessions coming to the website by year and month. At the top, the overall sessions is displayed

Traffic Sessions

Total Sessions
472871

month_description	2012	2013	2014	2015	Total
01 -- January		6401	14792	25191	**46384**
02 -- February		7181	16298	23804	**47283**
03 -- March	1903	6289	15710	14936	**38838**
04 -- April	3747	7991	17401		**29139**
05 -- May	3741	8394	17950		**30085**
06 -- June	3944	8362	17810		**30116**
07 -- July	4281	8919	19048		**32248**
08 -- August	6082	9129	18579		**33790**
09 -- September	6589	9616	19551		**35756**
10 -- October	8192	10800	21420		**40412**
11 -- November	13970	13961	26306		**54237**
12 -- December	10077	15816	28690		**54583**
Total	**62526**	**112859**	**233555**	**63931**	**472871**

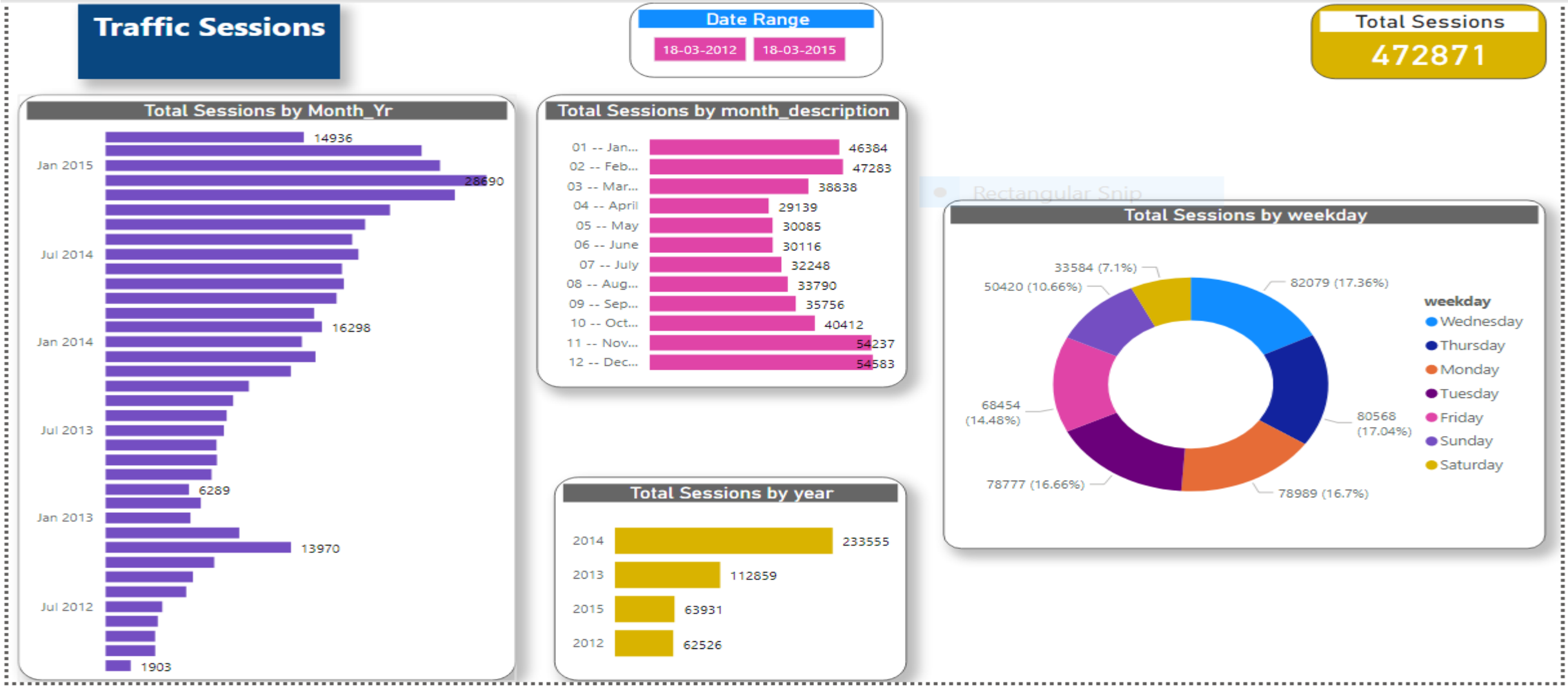
month_description	2012	2013	2014	2015	Total
01 -- January		6401	14792	25191	46384

The January month of various years have the Different count of sessions, it is increasing by year.

2013
6401
7181
6289
7991
8394
8362
8919
9129
9616
10800
13961
15816
112859

This is for the year 2013

This dashboard has a filter for selecting date range. Once selected one chart shows the session count distribution by Month Yr, another by month, another by year and the donut chart shows the distribution by weekdays. The date range selected is March 18th 2012 to March 18th 2015.



This dashboard displays charts for the date range of March 15th 2014 to March 18th of 2015

Traffic Sessions

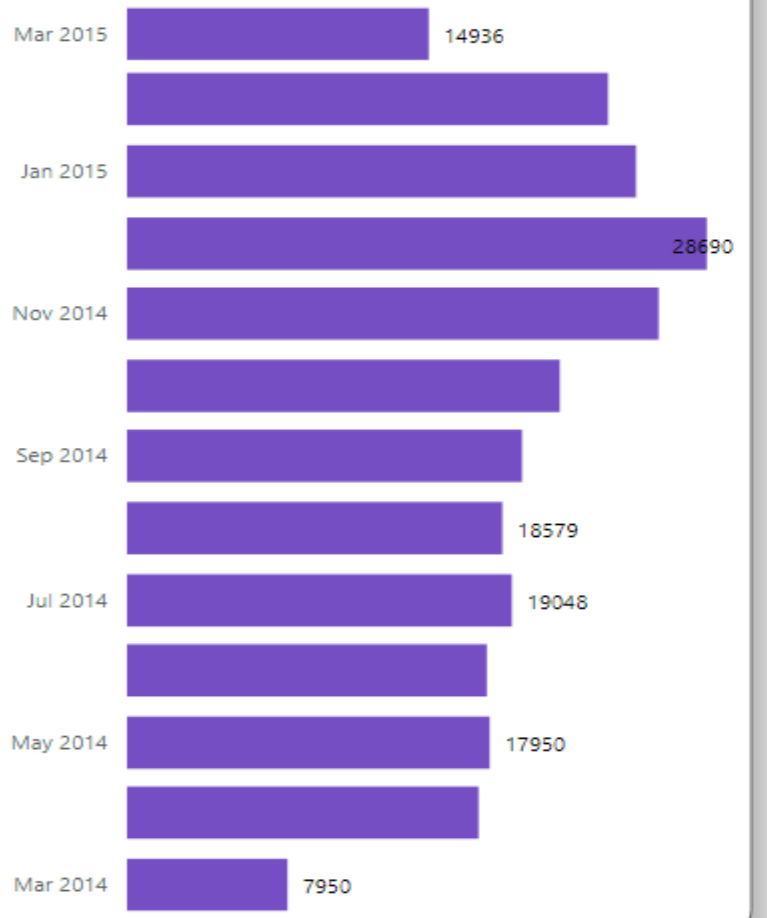
Date Range

15-03-2014 — 18-03-2015

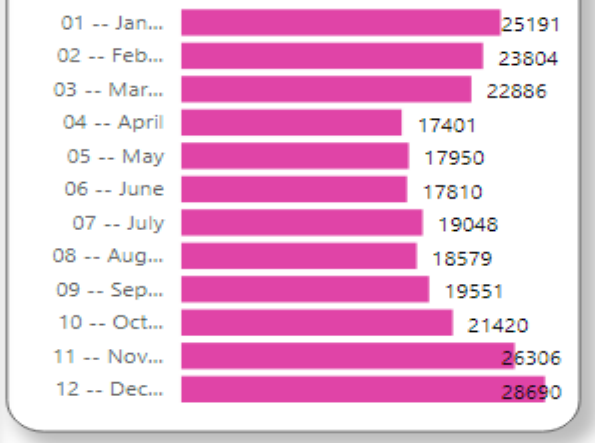
Total Sessions

258636

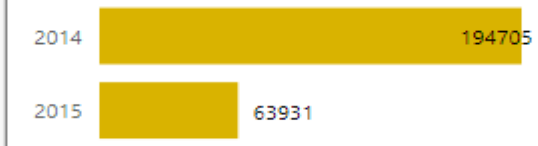
Total Sessions by Month_Yr



Total Sessions by month_description

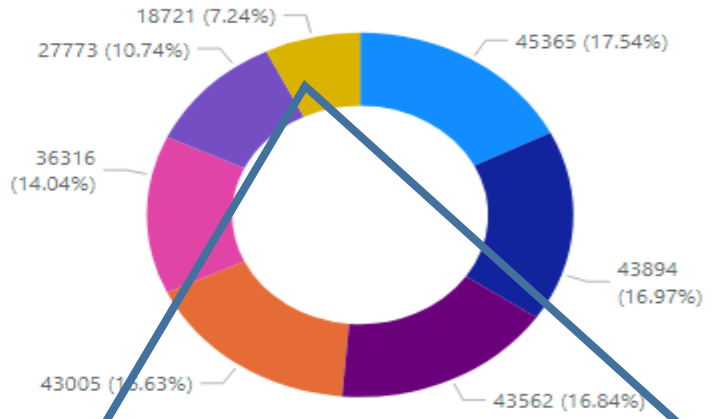


Total Sessions by year



The November, December, January, February Have high rate of sessions.

Total Sessions by weekday



Saturday has very low rate of sessions

This dashboard contains three filters, one is for utm source, another one for utm campaign and the last one is for device type. The utm sources contains gsearch and bsearch, the utm Campaign contains brand, nonbrand, desktop targeted, NULL and pilot. The device type contains desktop and mobile.

Traffic Split Sessions

UTM Source	UTM Campign	Device Type	Total Split Sessions
gsearch	nonbrand	All	282706

month_description	2012	2013	2014	2015	Total
01 -- January		3697	7464	13646	24807
02 -- February		4748	8243	13053	26044
03 -- March	1876	4095	8346	8300	22617
04 -- April	3516	5359	10341		19216
05 -- May	3298	5468	10637		19403
06 -- June	3424	5420	10478		19322
07 -- July	3693	5662	11149		20504
08 -- August	4646	5851	10419		20916
09 -- September	4254	6185	10607		21046
10 -- October	5207	6858	11373		23438
11 -- November	9227	9437	14738		33402
12 -- December	6497	10072	15422		31991
Total	45638	72852	129217	34999	282706

The values selected in the filters are gsearch and nonbrand

Traffic Split Sessions

UTM Source	UTM Campign	Device Type	Total Split Sessions
gsearch	nonbrand	desktop	195155

The filter values are
gsearch
nonbrand and
desktop.

month_description	2012	2013	2014	2015	Total
01 -- January		2799	4916	9541	17256
02 -- February		3612	5406	9145	18163
03 -- March	1144	3098	5468	5870	15580
04 -- April	2144	4061	6753		12958
05 -- May	2280	4220	6766		13266
06 -- June	2667	4105	6786		13558
07 -- July	2798	4335	7273		14406
08 -- August	3490	4147	6803		14440
09 -- September	3194	4004	6857		14055
10 -- October	3938	4498	7415		15851
11 -- November	6983	6089	10219		23291
12 -- December	4953	6596	10782		22331
Total	**33591**	**51564**	**85444**	**24556**	**195155**

The matrix shows the count of sessions
By year and month.
The total session count for the selected
Filter values is **195155.**

Traffic Split Sessions

UTM Source	UTM Campign	Device Type	Total Split Sessions
gsearch	nonbrand	mobile	87551

The filter values are
gsearch
nonbrand and
mobile.

month_description	2012	2013	2014	2015	Total
01 -- January		898	2548	4105	7551
02 -- February		1136	2837	3908	7881
03 -- March	732	997	2878	2430	7037
04 -- April	1372	1298	3588		6258
05 -- May	1018	1248	3871		6137
06 -- June	757	1315	3692		5764
07 -- July	895	1327	3876		6098
08 -- August	1156	1704	3616		6476
09 -- September	1060	2181	3750		6991
10 -- October	1269	2360	3958		7587
11 -- November	2244	3348	4519		10111
12 -- December	1544	3476	4640		9660
Total	**12047**	**21288**	**43773**	**10443**	**87551**

The matrix shows the count of sessions
By year and month.
The total session count for the selected
Filter values is **87551.**

The user can select the different combination of filter values to analyse the traffic of sessions

This dashboard shows the charts for distribution of session counts by Month yr, month, year and weekdays. The filters are synchronized with the previous page. Any value changed in the filter will be reflected on both Pages.

Traffic Split Sessions

Date Range: 18-03-2012 – 18-03-2015

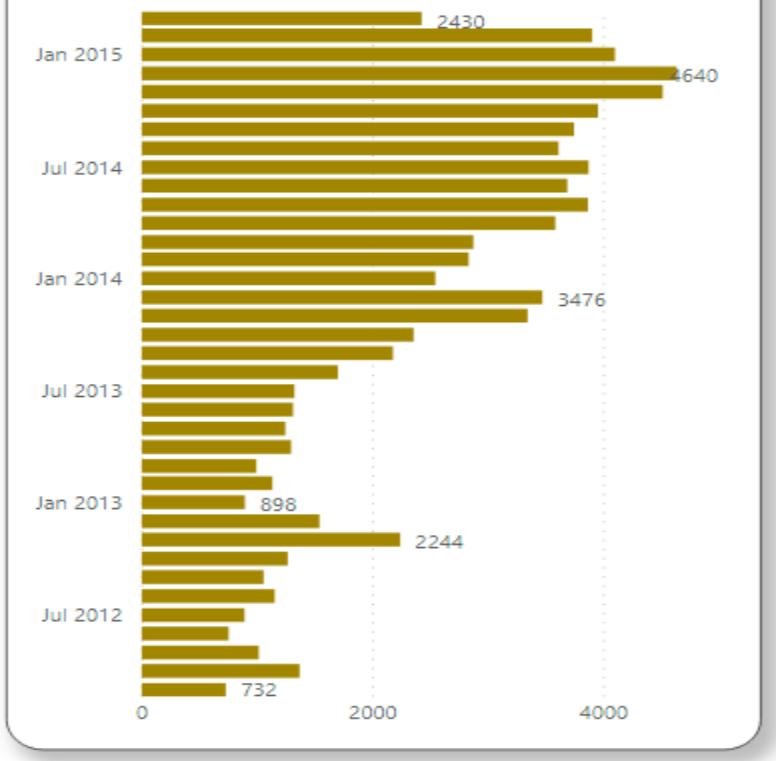
UTM Source: gsearch

UTM Campign: nonbrand

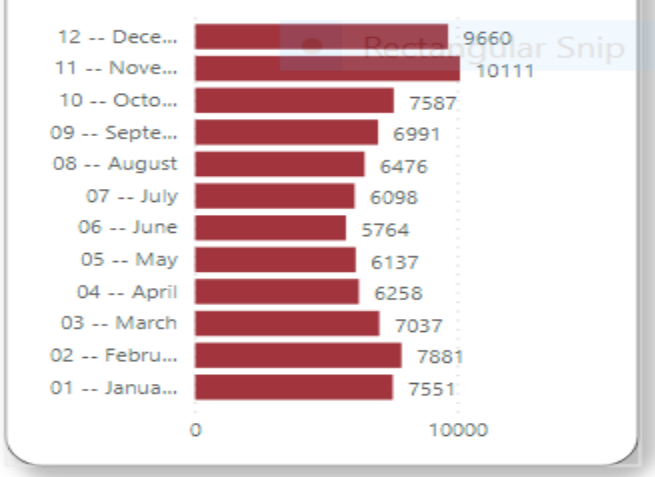
Device Type: mobile

Total Split Sessions: 87551

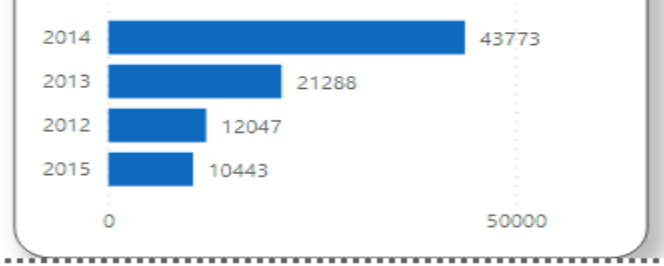
Total Split Sessions by Month_Yr



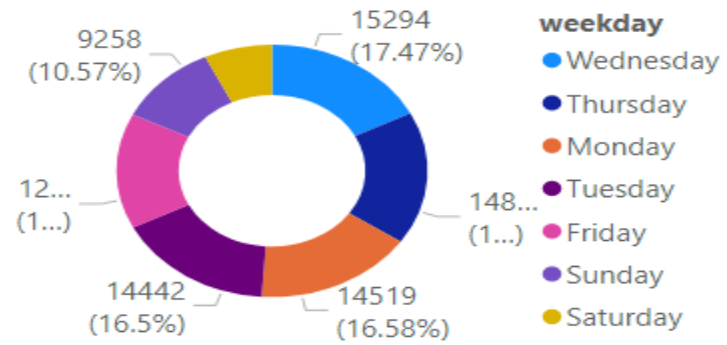
Total Split Sessions by month_description



Total Split Sessions by year



Total Split Sessions by weekday



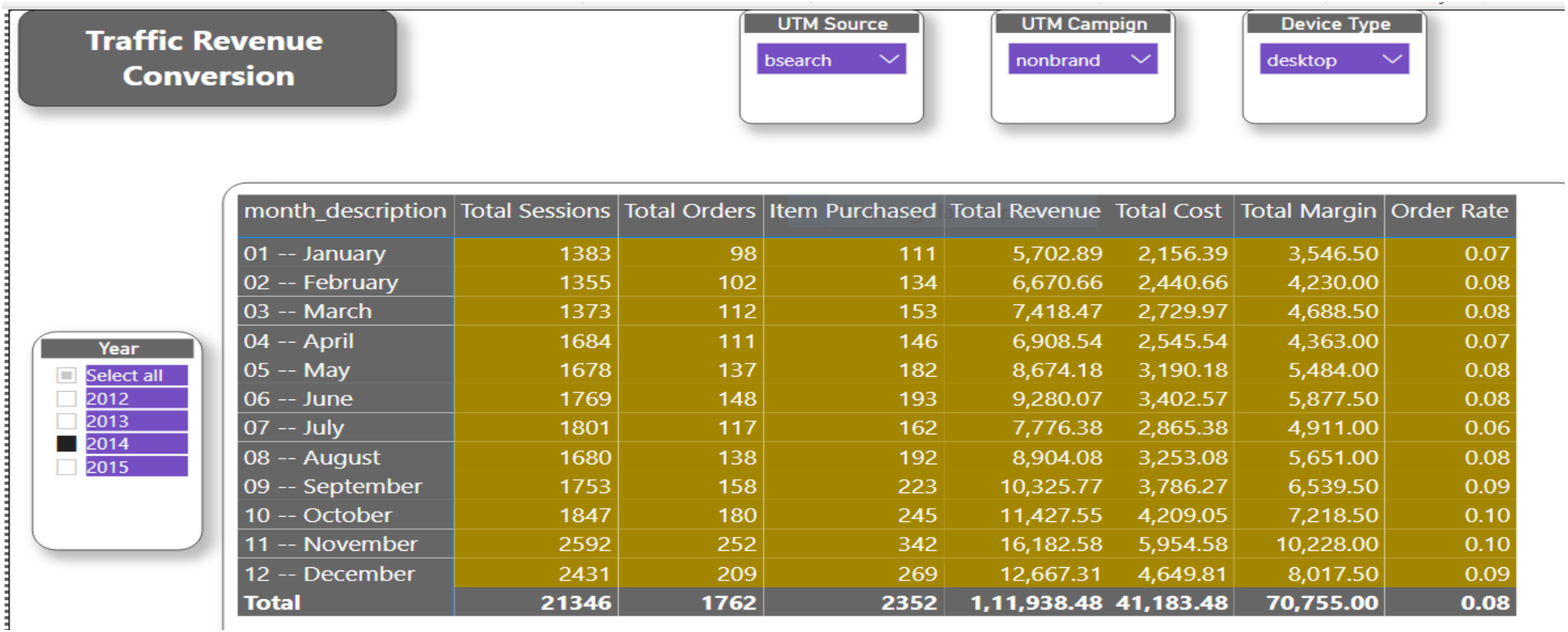
Revenue Conversion

To know how the sessions are converted in to revenue is important for the business.

It is also important to know which source and campaign help generate more revenue.

If revenue generated by any source is not up to the expectations, then the business can device strategies to improve the revenue converted by the sessions coming through that source.

This report contains four filters. Year, UTM Source, UTM Campaign and Device Type. The selected values are 2014, Bsearch, nonbrand and desktop. The matrix shows values for the measurements total sessions, total orders, total Item purchased, total revenue, total cost, total margin and order rate by year and month.
The order rate is calculated as total orders divided by total sessions.



month_description	Total Sessions	Total Orders	Item Purchased	Total Revenue	Total Cost	Total Margin	Order Rate
01 -- January	1383	98	111	5,702.89	2,156.39	3,546.50	0.07
02 -- February	1355	102	134	6,670.66	2,440.66	4,230.00	0.08
03 -- March	1373	112	153	7,418.47	2,729.97	4,688.50	0.08
04 -- April	1684	111	146	6,908.54	2,545.54	4,363.00	0.07
05 -- May	1678	137	182	8,674.18	3,190.18	5,484.00	0.08
06 -- June	1769	148	193	9,280.07	3,402.57	5,877.50	0.08
07 -- July	1801	117	162	7,776.38	2,865.38	4,911.00	0.06
08 -- August	1680	138	192	8,904.08	3,253.08	5,651.00	0.08
09 -- September	1753	158	223	10,325.77	3,786.27	6,539.50	0.09
10 -- October	1847	180	245	11,427.55	4,209.05	7,218.50	0.10
11 -- November	2592	252	342	16,182.58	5,954.58	10,228.00	0.10
12 -- December	2431	209	269	12,667.31	4,649.81	8,017.50	0.09
Total	**21346**	**1762**	**2352**	**1,11,938.48**	**41,183.48**	**70,755.00**	**0.08**

This is the report when the filter values are selected as 2013, gsearch, brand and desktop

Traffic Revenue Conversion

UTM Source: gsearch

UTM Campign: brand

Device Type: desktop

Year
- Select all
- 2012
- 2013
- 2014
- 2015

month_description	Total Sessions	Total Orders	Item Purchased	Total Revenue	Total Cost	Total Margin	Order Rate
01 -- January	307	26	26	1,339.74	518.74	821.00	0.08
02 -- February	226	33	33	1,769.67	679.17	1,090.50	0.15
03 -- March	206	14	14	719.86	278.86	441.00	0.07
04 -- April	258	28	28	1,459.72	563.72	896.00	0.11
05 -- May	275	22	22	1,149.78	443.78	706.00	0.08
06 -- June	296	34	34	1,759.66	680.66	1,079.00	0.11
07 -- July	365	27	27	1,409.73	544.23	865.50	0.07
08 -- August	346	37	37	1,929.63	745.13	1,184.50	0.11
09 -- September	356	37	37	1,909.63	739.13	1,170.50	0.10
10 -- October	421	45	46	2,369.54	917.54	1,452.00	0.11
11 -- November	473	48	49	2,489.51	967.01	1,522.50	0.10
12 -- December	616	70	73	3,735.27	1,406.77	2,328.50	0.11
Total	**4145**	**421**	**426**	**22,041.74**	**8,484.74**	**13,557.00**	**0.10**

This is the running chart for sessions, orders and order rate for the year 2013 with the selected Values of gsearch, brand and desktop.

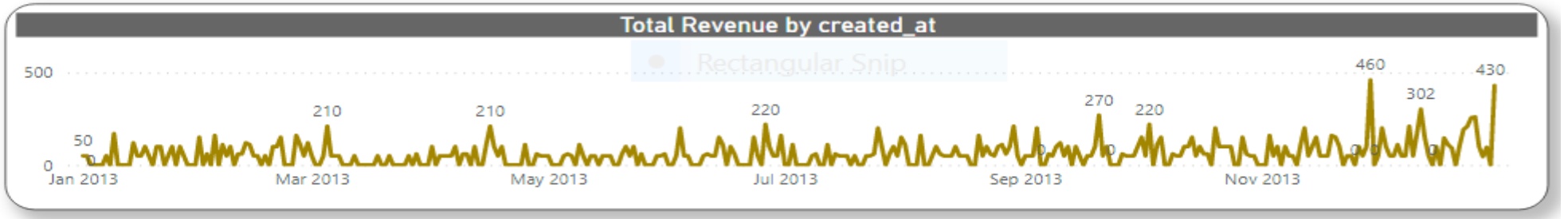
This is for Total Revenue, Total Cost and Total Margin

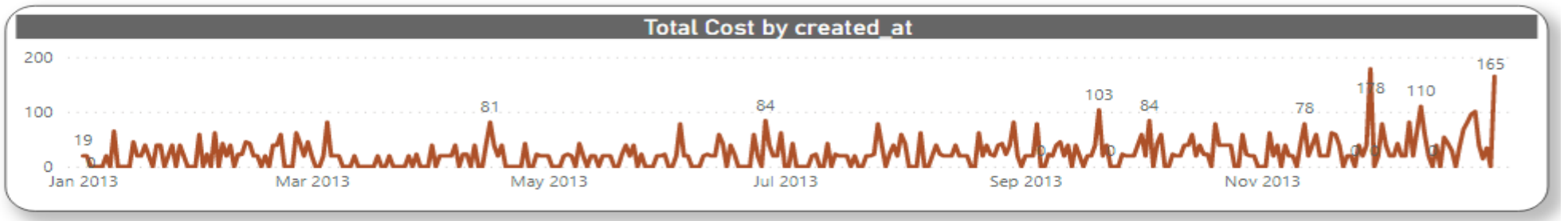
Traffic Revenue Conversion

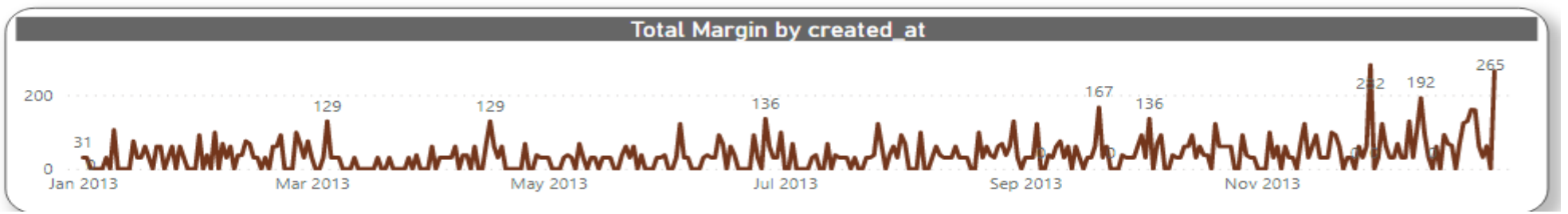
Month	UTM Source	UTM Campign	Device Type
All	gsearch	brand	desktop

Year

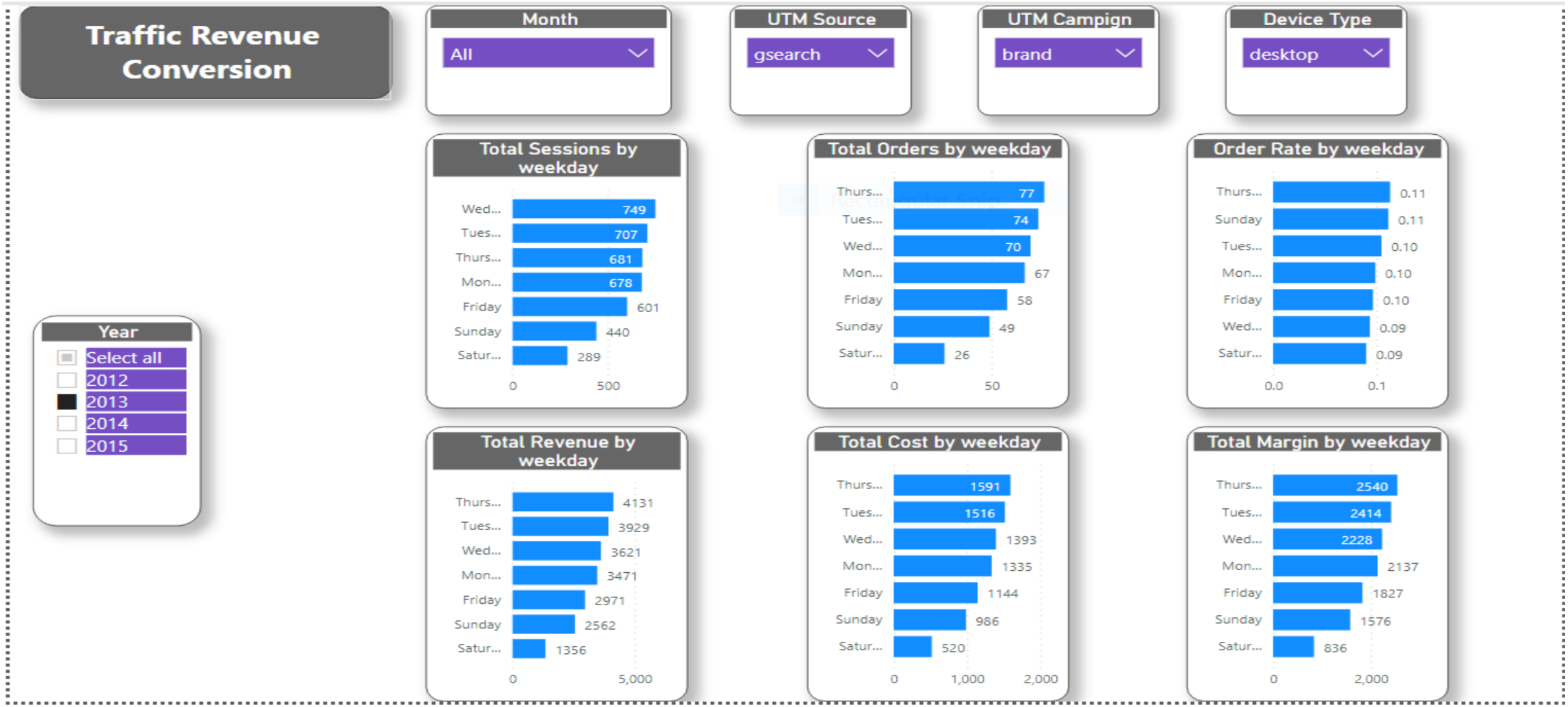
- [ ] Select all
- [ ] 2012
- [x] 2013
- [ ] 2014
- [ ] 2015







This is for analysing the various measures by weekdays. Which week day brings high rate of sessions and which Has high rate of revenue generation (Thursday). We now know that the filters for all four pages are synchronized.



Thank You