

19162121031

SMIT R PATEL

SEM 5

PRACTICAL 16

BIG DATA AND ANALYTICS

AIM- To understand the working of SPSS Modeller in Watson Studio.

Exercise: Vodafone wants to analyse the factors as to why their customers are disassociating with their company and churning out. They have a few data files in order to help understand the same. Use a tool in the IBM Cloud in order to create a model to get insights and obtain graphs to understand the results better.

Tasks:

1. Upload the data files in your Watson Studio Instance.

The screenshot displays the IBM Watson Studio web interface. The top navigation bar includes the IBM Watson Studio logo, a search bar, and user account information for Smit Patel. The main content area is divided into tabs: Overview, Assets (selected), Environments, Jobs, Access control, and Settings. Under the Assets tab, there is a search bar and a section titled 'Data assets' showing 0 assets selected. A table lists five CSV data assets, all created by Smit Patel on October 7, 2021, at 12:57 PM. The table columns are Name, Type, Created by, and Last modified. A right-hand sidebar shows a 'Data' panel with options to Load, Files, or Catalog, and a prompt to drop files or browse for upload. The bottom of the screen shows a Windows taskbar with the time 1:03 PM on 10/7/2021 and an 'Activate Windows' watermark.

Name	Type	Created by	Last modified
telco x data.csv	Data Asset	Smit Patel	Oct 07, 2021, 12:57 PM
telco x call data q2.csv	Data Asset	Smit Patel	Oct 07, 2021, 12:57 PM
telco x call data q1.csv	Data Asset	Smit Patel	Oct 07, 2021, 12:57 PM
telco x customer data.csv	Data Asset	Smit Patel	Oct 07, 2021, 12:57 PM
telco x products.csv	Data Asset	Smit Patel	Oct 07, 2021, 12:57 PM

2. Understand basics of SPSS and a modeller flow.

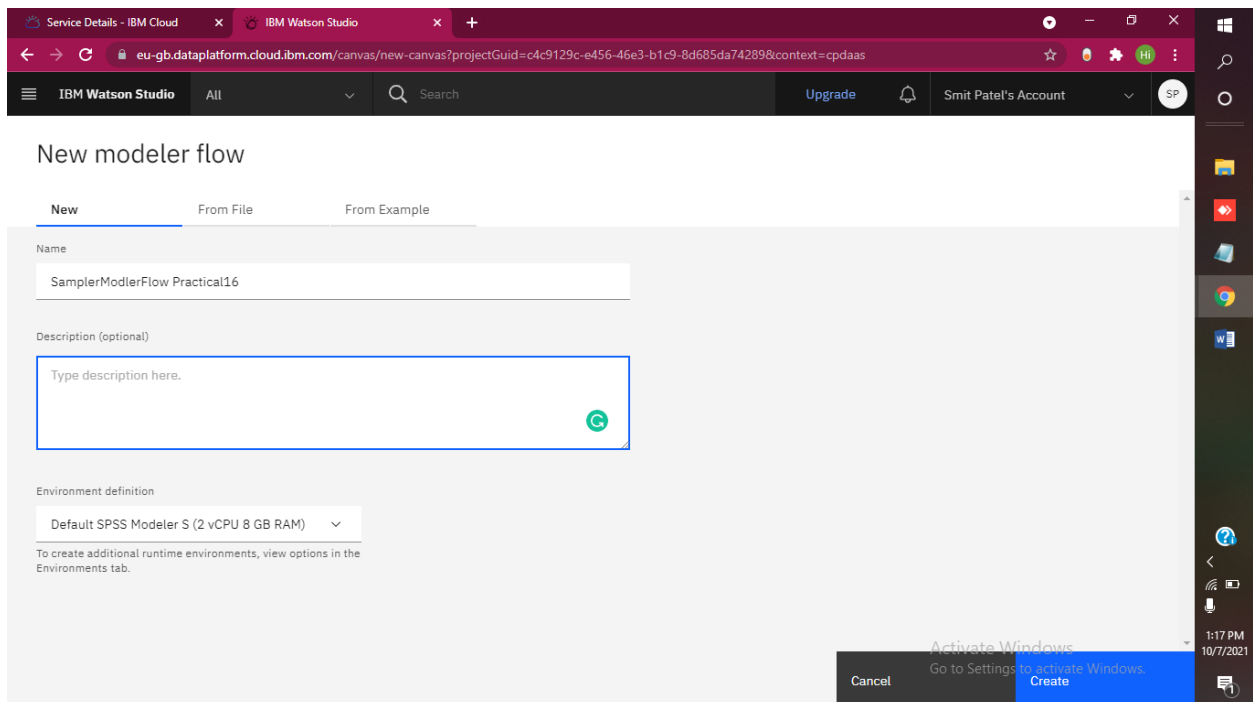
→ SPSS stands for **Statistical Package for the Social Sciences**

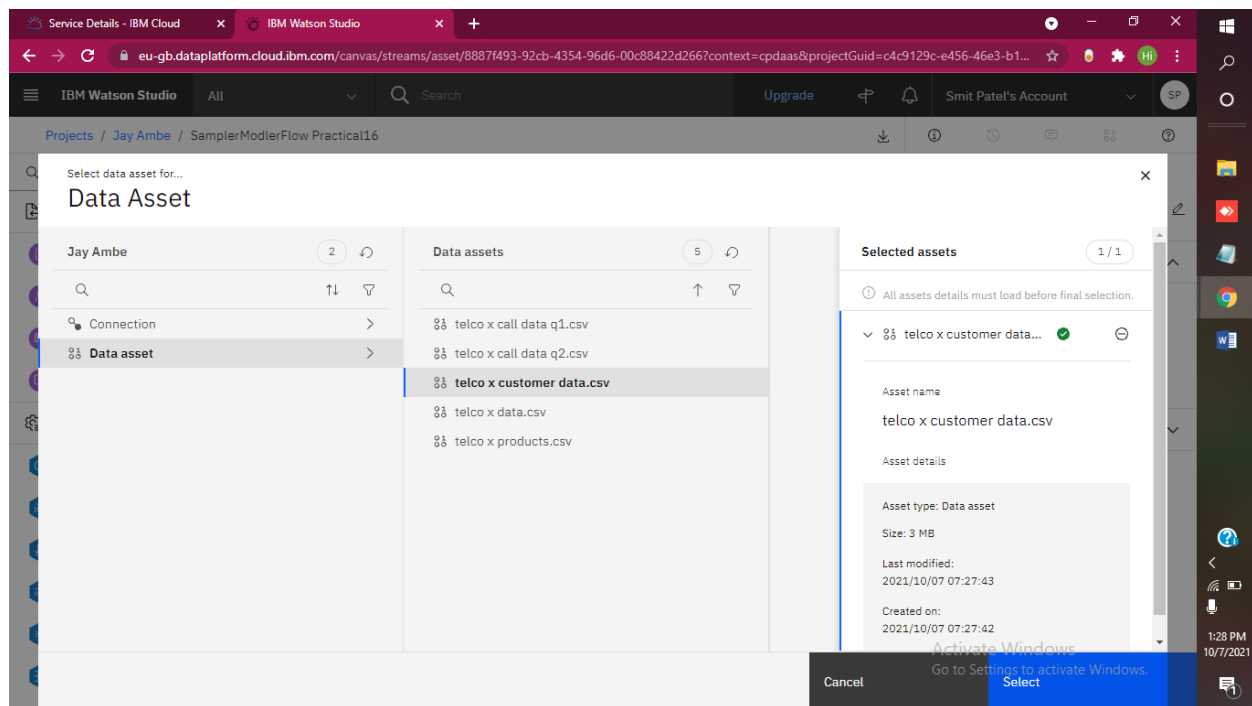
→ IBM Watson SPSS Modeler flows in Watson Studio provide an interactive environment for quickly building machine learning pipelines that **flow data from ingestion to transformation to model building and evaluation**, without needing any code.

→ SPSS Modeler provides data mining and text analysis software, while SPSS Statistics is an integrated family of products. Both tools enable users **to build predictive models** and execute other analytics tasks. Modeler can apply different processes and algorithms to help the user discover information hidden in the data

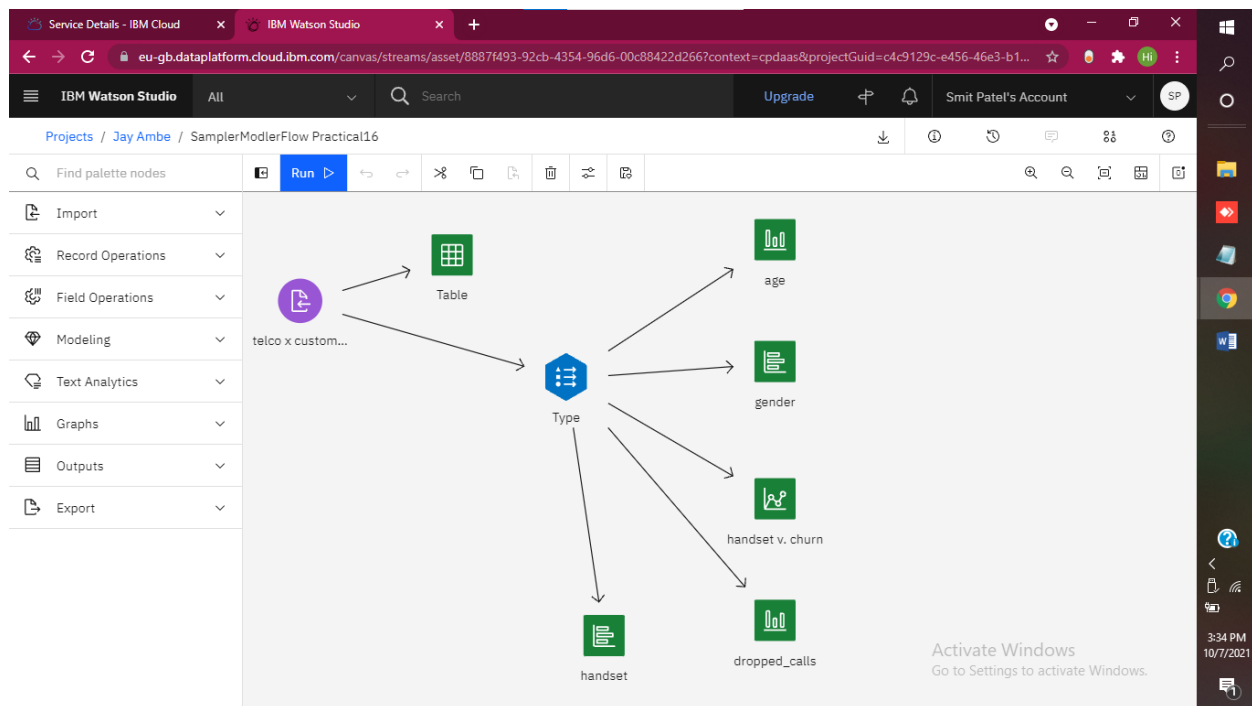
→ **IBM SPSS Statistics** (2010) The new official name is now IBM SPSS Statistics and is part of IBM's analytics portfolio

3. Create an appropriate SPSS modeller flow to obtain the result.





4. Add graph nodes to depict results in the form of Graphs.



Service Details - IBM Cloud

IBM Watson Studio

+

eu-gb.dataplatform.cloud.ibm.com/canvas/streams/asset/8887f493-92cb-4354-96d6-00c88422d266?context=cpdaas&projectGuid=c4c9129c-e456-46e3-b1...

IBM Watson Studio

All

Search

Upgrade

Smit Patel's Account

SP

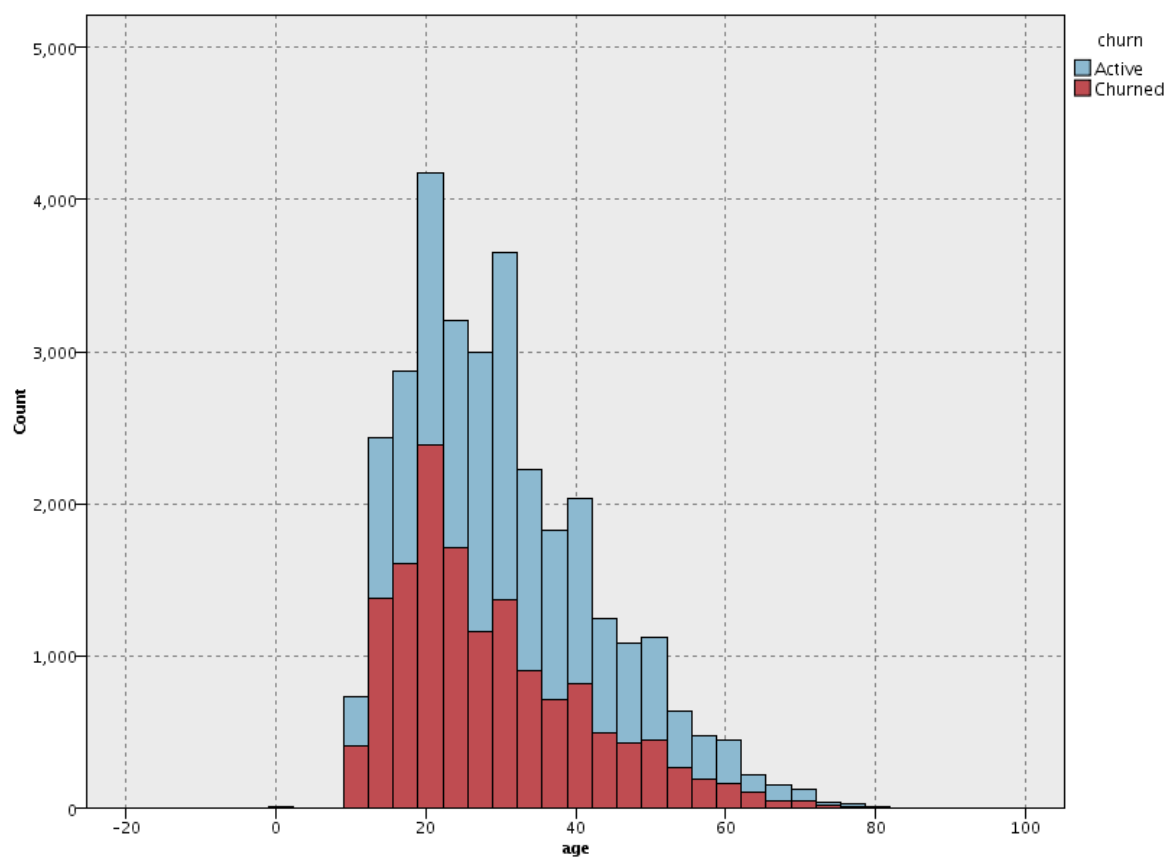
Projects / Jay Ambe / SamplerModlerFlow Practical16

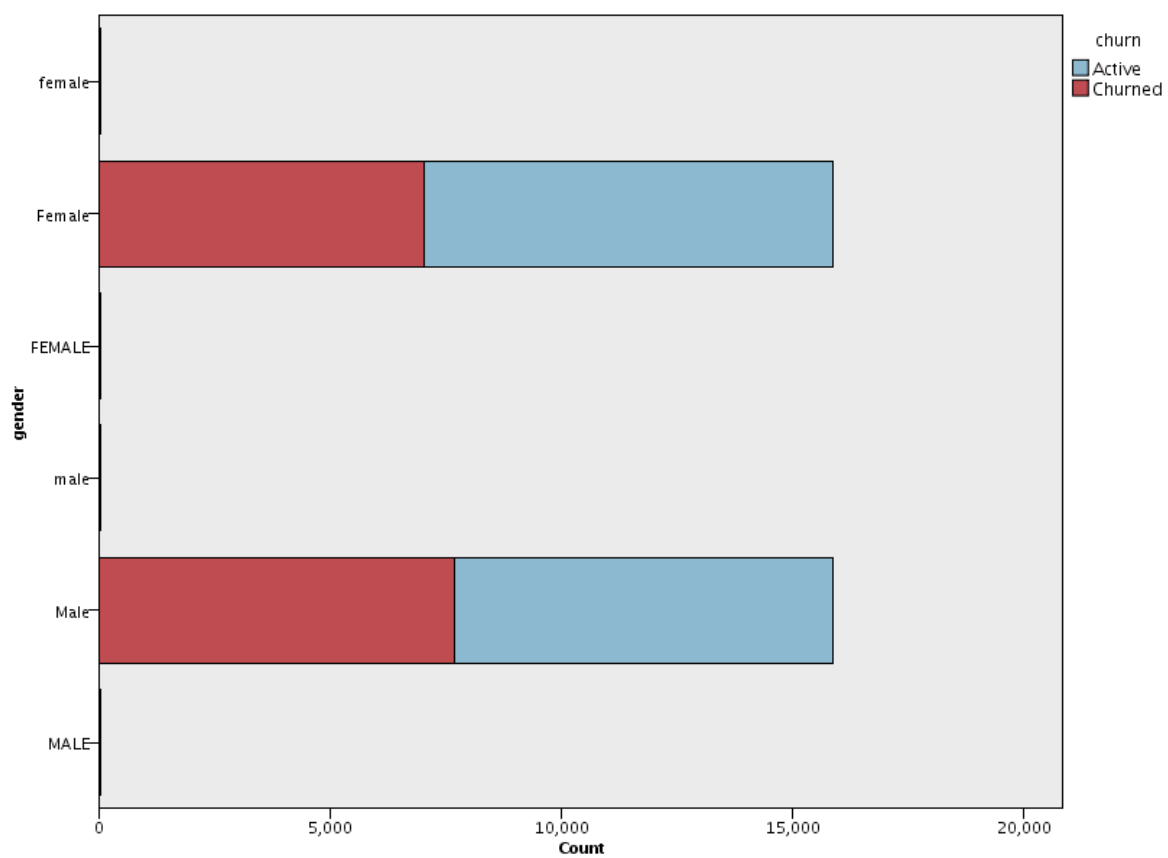
View Output: Table (12 fields, 31,781 records) #2

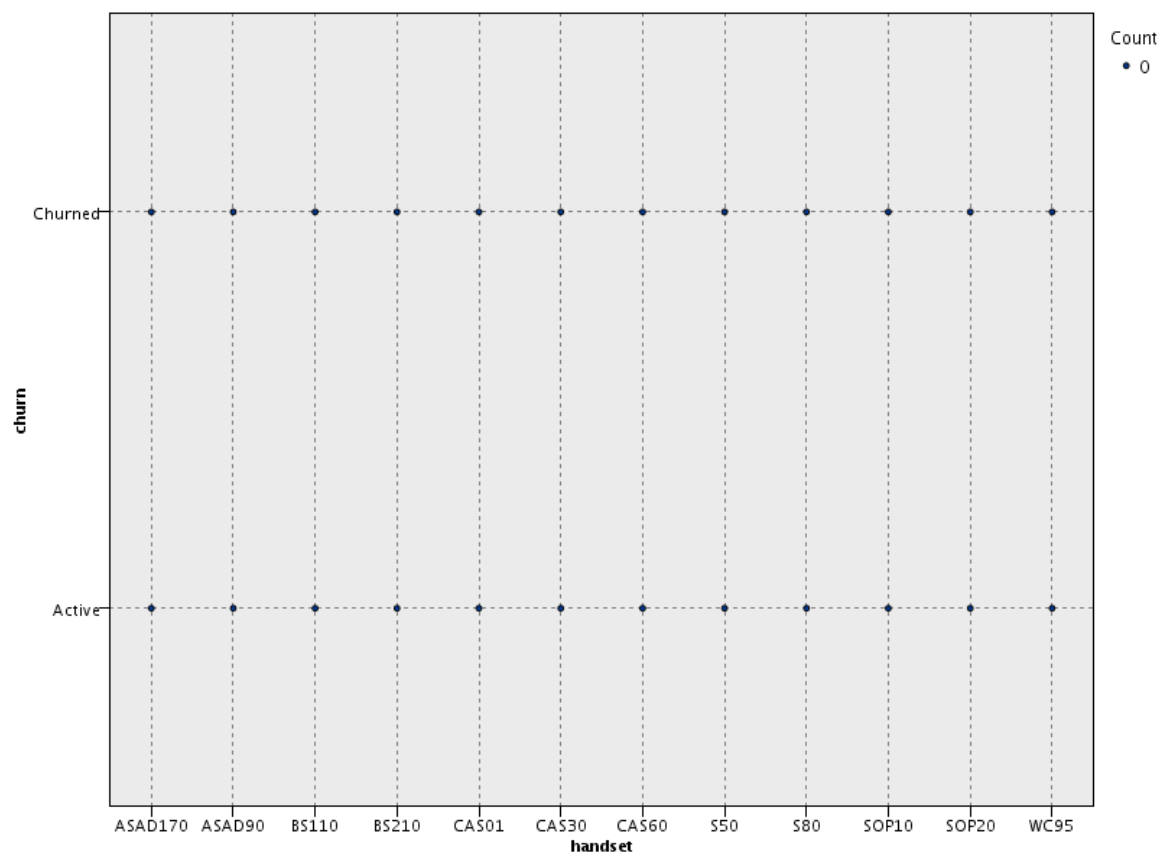
customer_id	gender	age	postalcode	region	connect_date	end_date	dropped_calls	pay method	tariff	handset	churn
K100010	Male	46.000	6253.000	3.000	2006-07-18	2010-05-10	1.000	Pre Pay	CAT 50	SOP10	Churned
K100020	Male	27.000	4121.000	2.000	2004-09-18	2007-02-07	0.000	Pre Pay	CAT 50	SOP10	Churned
K100030	Male	39.000	3870.000	2.000	2003-08-23	2006-08-13	2.000	Pre Pay	CAT 50	SOP20	Churned
K100040	Male	28.000	8322.000	4.000	2005-08-18	2006-06-04	2.000	Pre Pay	CAT 50	SOP10	Churned
K100050	Male	47.000	2614.000	2.000	2005-08-11	2006-08-13	0.000	Pre Pay	CAT 50	SOP10	Churned
K100060	Male	29.000	1891.000	1.000	2004-07-06	2005-03-13	1.000	Pre Pay	CAT 50	SOP10	Churned
K100070	Male	38.000	1741.000	1.000	2006-10-14	2007-01-01	1.000	Pre Pay	CAT 50	SOP20	Churned
K100080	Male	27.000	9770.000	4.000	2005-01-20	2009-08-27	0.000	Pre Pay	CAT 50	SOP20	Churned
K100090	Male	38.000	8967.000	4.000	2003-05-31	2004-05-19	1.000	Pre Pay	CAT 50	SOP20	Churned

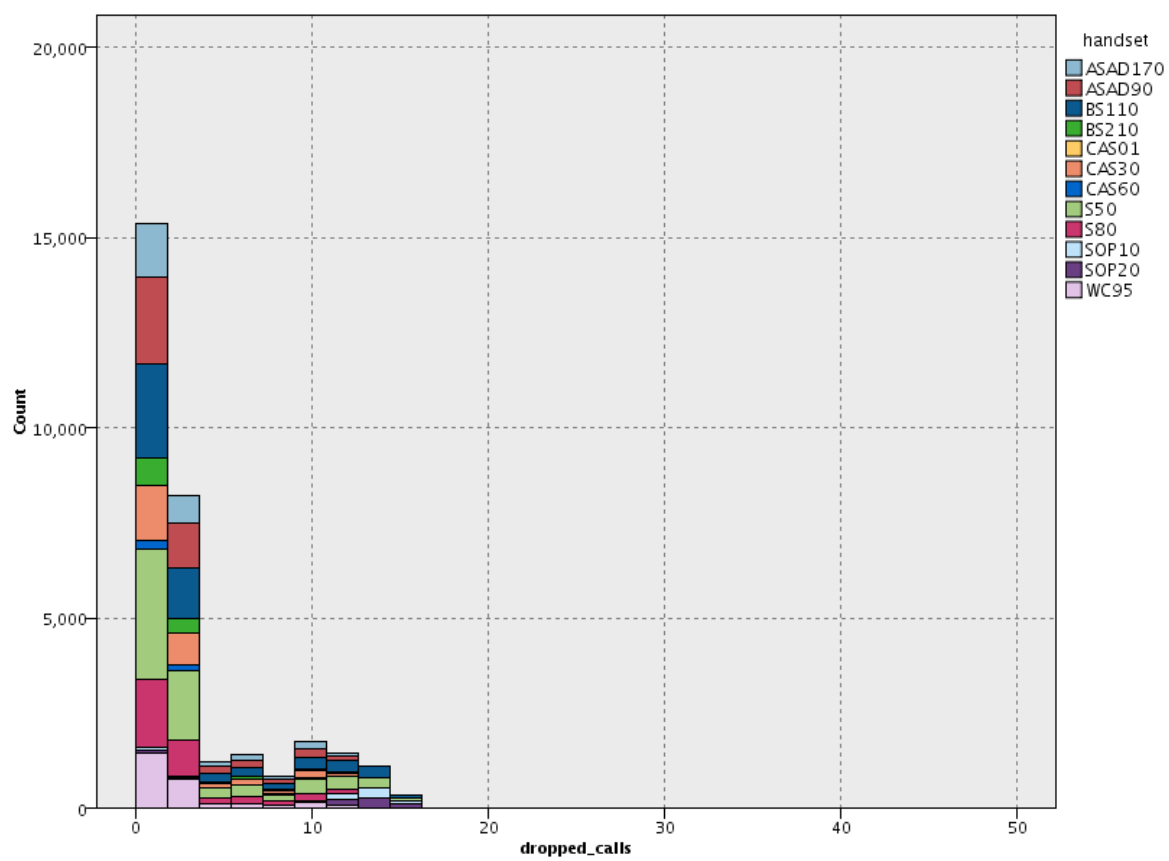
Activate Windows
Go to Settings to activate Windows.

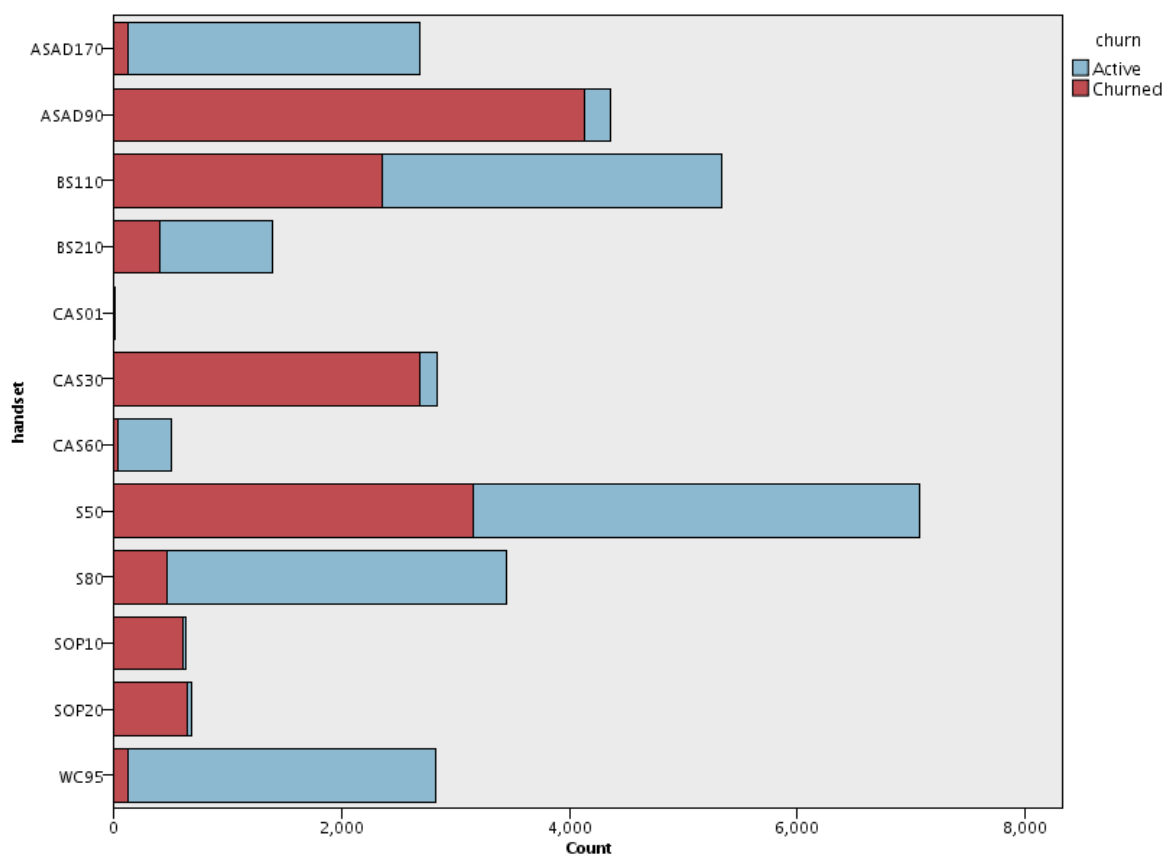
3:35 PM
10/7/2021





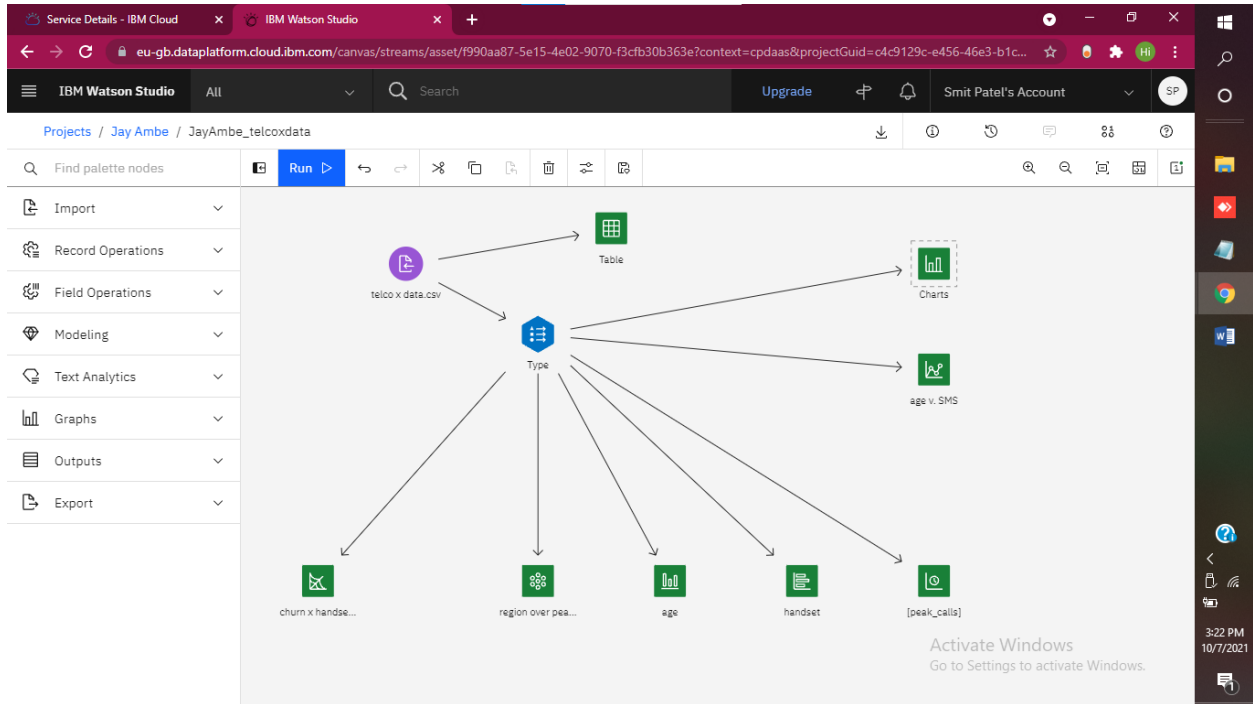






Extra :-

Now We get telco x data.csv file for practice more in Watson studio



View Output: (54 fields, 5,000 records) #1

customer_id	gender	age	postalcode	region	tariff	connect_date	end_date	dropped_calls	pay method	handset	churn	peak_
K338270	Female	20.000	6599.000	3.000	CAT 100	2006-09-08		1.000	Pre Pay	WC95	Active	240.00
K342660	Male	21.000	1635.000	1.000	CAT 100	2006-04-09		7.000	Pre Pay	S50	Active	281.00
K342650	Male	17.000	2149.000	1.000	CAT 100	2003-10-11		6.000	Pre Pay	S50	Active	151.00
K342640	Male	20.000	7788.000	4.000	CAT 100	2004-11-24		1.000	Pre Pay	S50	Active	340.00
K342630	Male	46.000	5223.000	3.000	CAT 100	2004-01-16		2.000	Pre Pay	CAS30	Active	193.00
K342620	Female	24.000	9143.000	4.000	CAT 100	2005-06-26		3.000	Pre Pay	BS110	Active	184.00
K342610	Male	32.000	4223.000	2.000	CAT 100	2003-07-26		1.000	Pre Pay	S50	Active	298.00
K342600	Female	51.000	6165.000	3.000	CAT 100	2005-09-26	2007-06-15	0.000	Pre Pay	CAS30	Churned	287.00
K342590	Male	23.000	7210.000	3.000	CAT 100	2003-10-24		1.000	Pre Pay	BS110	Active	215.00

Activate Windows
Go to Settings to activate Windows.

