# BUSINESS INTELLIGENCE PROJECT CASE STUDY

NAME : SYED SALMAN HAIDER (09419)

TOPIC : **Energy and Mining Statistics**

Reference : http://www.pbs.gov.pk/content/social-statistics

Instructor : Dr. Tariq MAHMOOD

# ABOUT DATA :

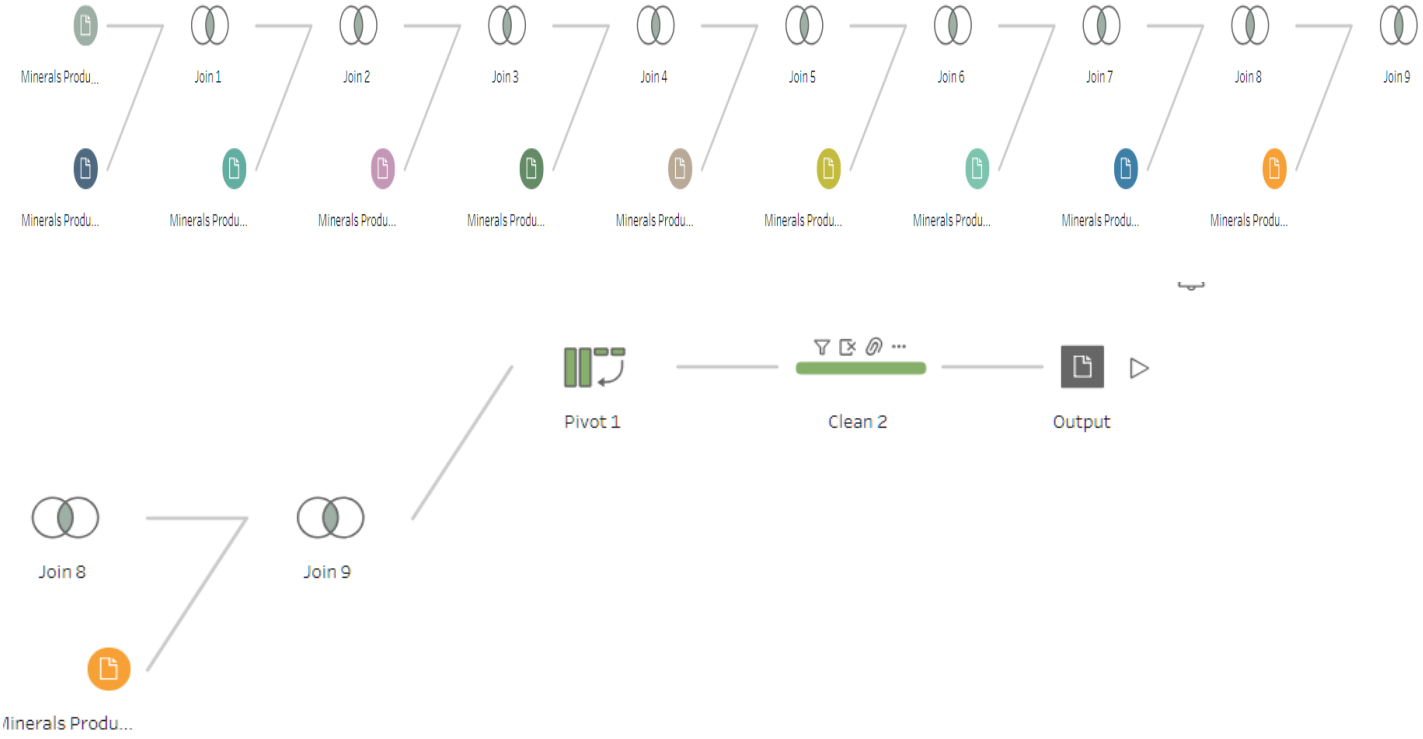
Monthly Mineral production data in respect of four provinces is received from provincial Directorate of Mines and Mineral and Directorate General of Petroleum Concessions, ministry of Petroleum and natural Resources. This monthly data is then aggregated to form the annual figure. Thereafter, Monthly and Yearly mineral production data in respect of 40 selected mineral items are also published and also supplied to various National and international agencies.

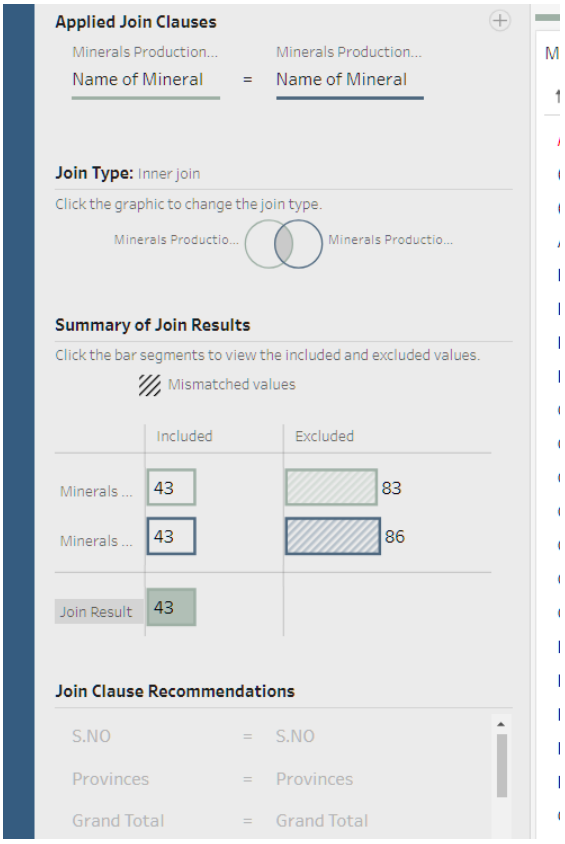
Pakistan Bureau of Statistics has been computing annual series of Quantum Indices of Mining Production based on weights derived from census value added of mining sector, which was designed to measure changes in physical out-put of mining indices. Three series of mining indices have so for been computed by taking 1975-76, 1980-81 and 1999-2000 as base. Now the work on the development of indices with the base 2005-06 is in progress.

# DATA CLEANING / DATA PRE-PROSSESING

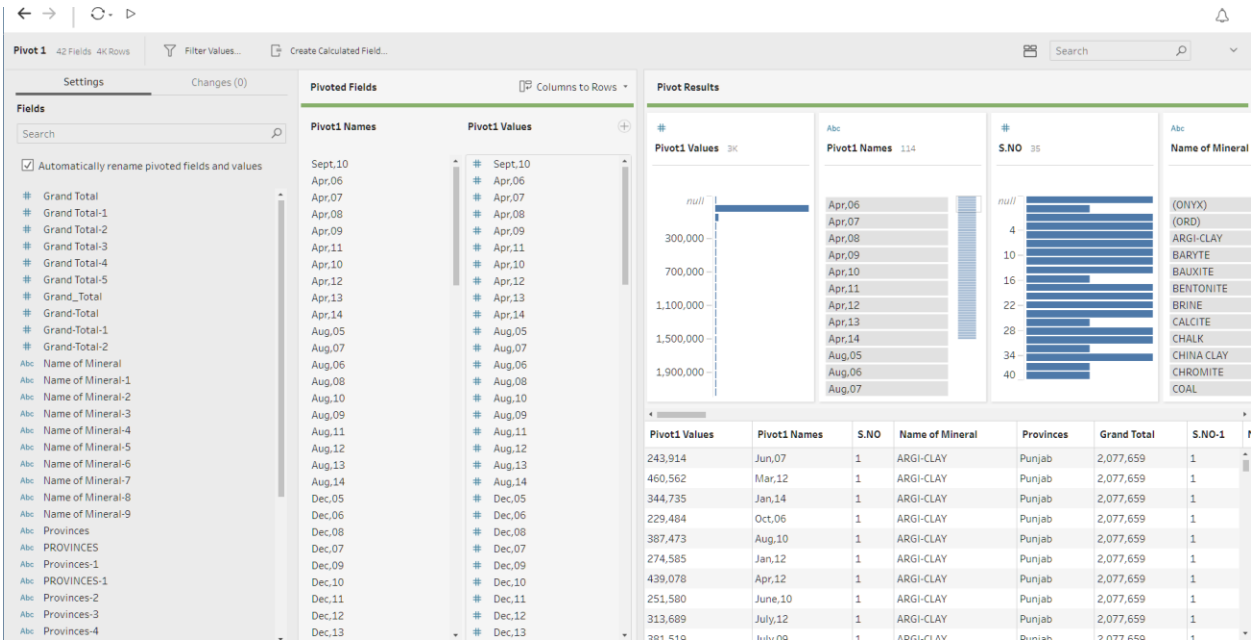
There were 10 Minerals Production data pdf files. I first converted them to excel then each excel file had different sheets which were combine and later all excel files were joined using tableau then I cleaned the data set using tableau prep so it can be processed further.

Below are the snapshots:

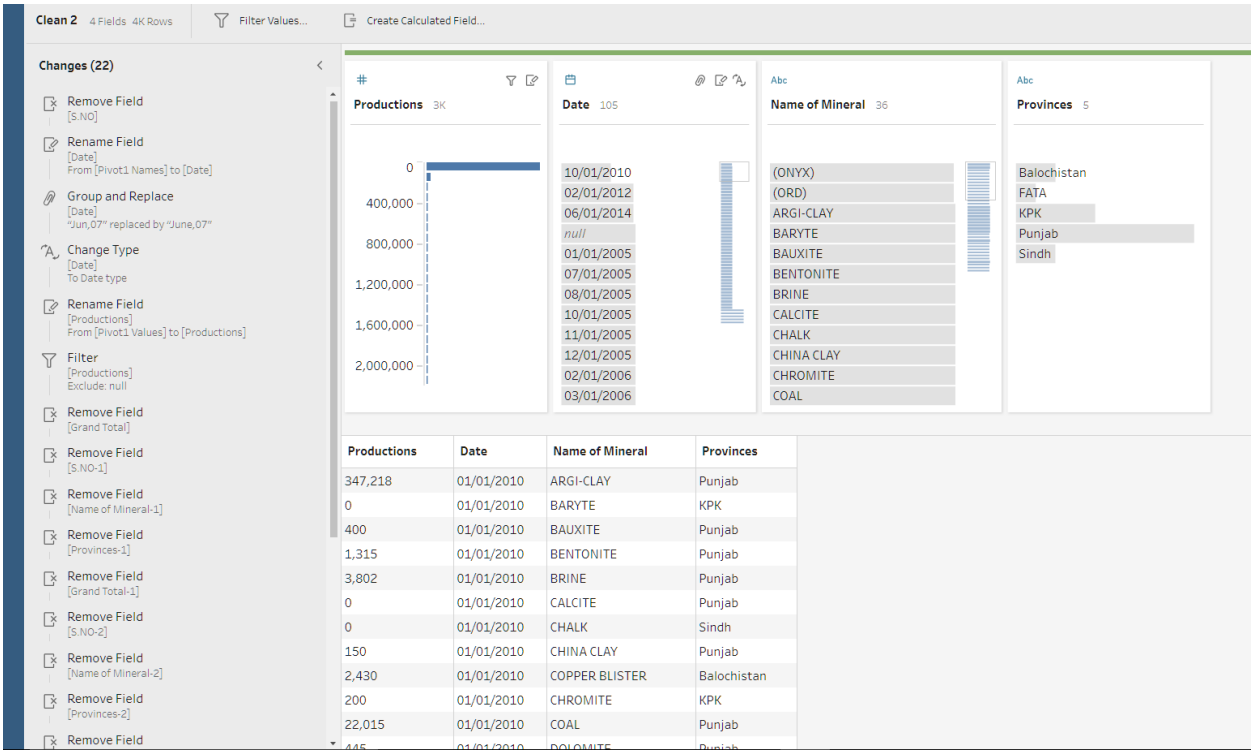




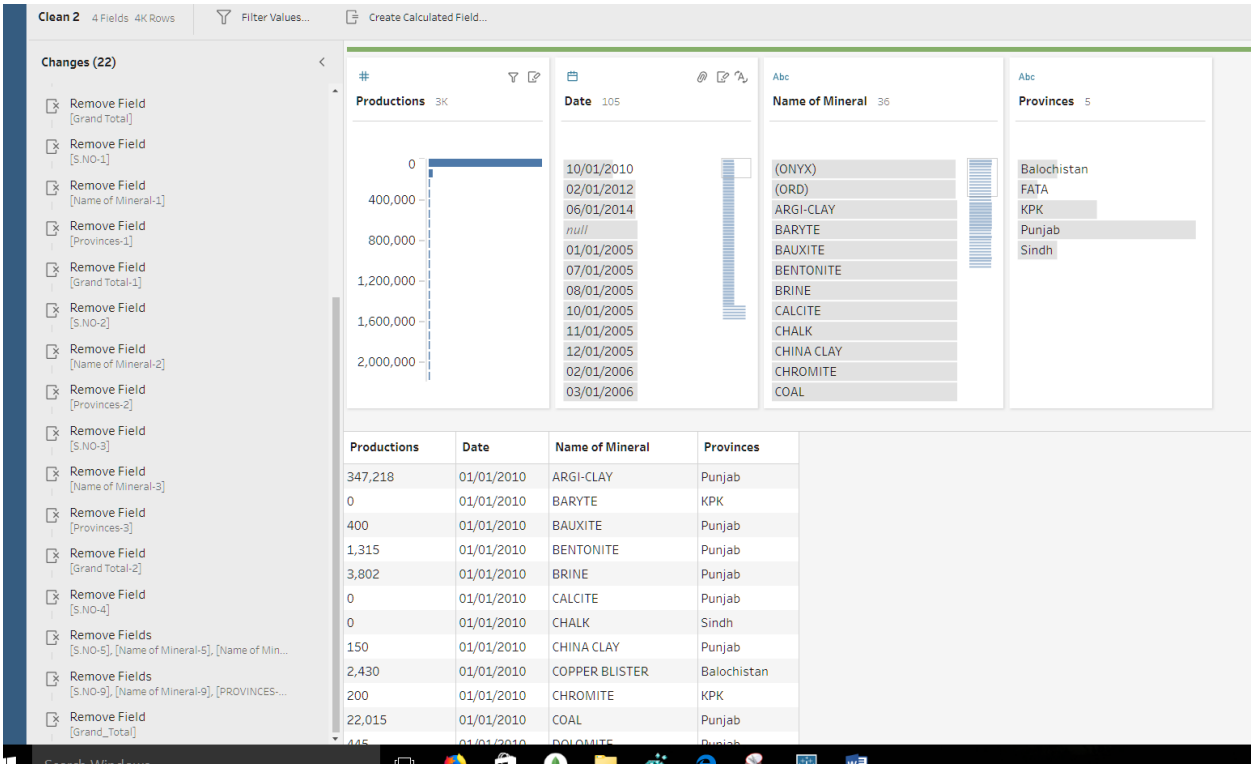

- I used inner join so I can have matching rows with the clause “Name of mineral” as it was the unique value for each record.




- As the date was given In columns which is not the right way to put the data and harder for analyses , so I converted them into rows by using pivot option



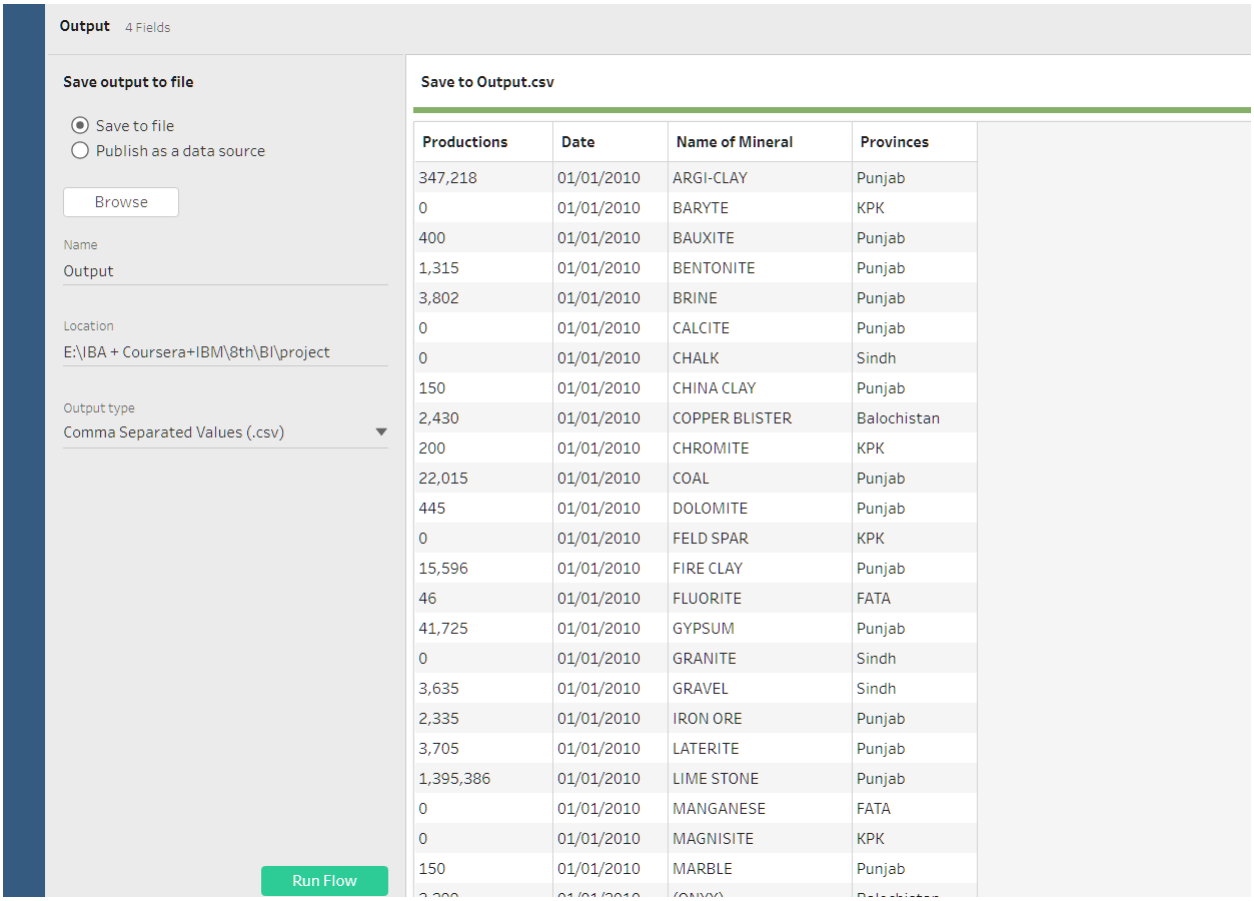

| Productions | Date | Name of Mineral | Provinces |
|---|---|---|---|
| 347,218 | 01/01/2010 | ARGI-CLAY | Punjab |
| 0 | 01/01/2010 | BARYTE | KPK |
| 400 | 01/01/2010 | BAUXITE | Punjab |
| 1,315 | 01/01/2010 | BENTONITE | Punjab |
| 3,802 | 01/01/2010 | BRINE | Punjab |
| 0 | 01/01/2010 | CALCITE | Punjab |
| 0 | 01/01/2010 | CHALK | Sindh |
| 150 | 01/01/2010 | CHINA CLAY | Punjab |
| 2,430 | 01/01/2010 | COPPER BLISTER | Balochistan |
| 200 | 01/01/2010 | CHROMITE | KPK |
| 22,015 | 01/01/2010 | COAL | Punjab |




| Productions | Date | Name of Mineral | Provinces |
|---|---|---|---|
| 347,218 | 01/01/2010 | ARGI-CLAY | Punjab |
| 0 | 01/01/2010 | BARYTE | KPK |
| 400 | 01/01/2010 | BAUXITE | Punjab |
| 1,315 | 01/01/2010 | BENTONITE | Punjab |
| 3,802 | 01/01/2010 | BRINE | Punjab |
| 0 | 01/01/2010 | CALCITE | Punjab |
| 0 | 01/01/2010 | CHALK | Sindh |
| 150 | 01/01/2010 | CHINA CLAY | Punjab |
| 2,430 | 01/01/2010 | COPPER BLISTER | Balochistan |
| 200 | 01/01/2010 | CHROMITE | KPK |
| 22,015 | 01/01/2010 | COAL | Punjab |

- The cleaning steps or all the changes I made in data are shown above in the changes area. There were irrelevant names like for June's month somewhere it was written jun and somewhere it was written June so I fixed that, there were few missing values and the date was not in string

format and I converted it to the data format, there were few irrelevant columns so I removed them as they were not needed in analysis and all other changes are listed above.



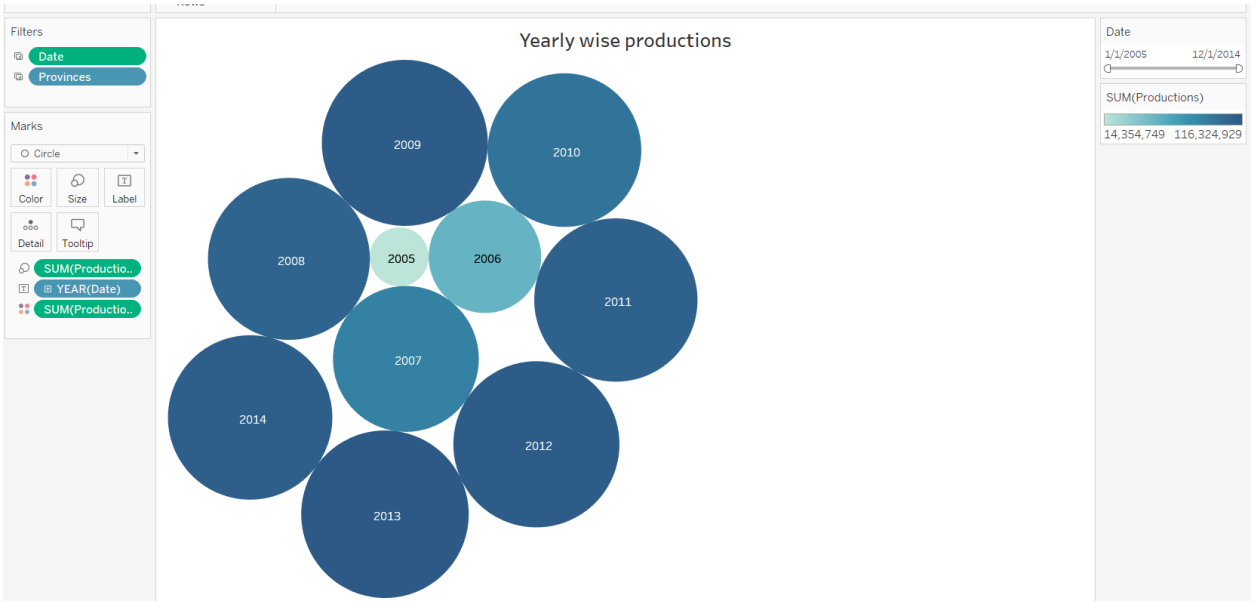

| Productions | Date | Name of Mineral | Provinces |
|---|---|---|---|
| 347,218 | 01/01/2010 | ARGI-CLAY | Punjab |
| 0 | 01/01/2010 | BARYTE | KPK |
| 400 | 01/01/2010 | BAUXITE | Punjab |
| 1,315 | 01/01/2010 | BENTONITE | Punjab |
| 3,802 | 01/01/2010 | BRINE | Punjab |
| 0 | 01/01/2010 | CALCITE | Punjab |
| 0 | 01/01/2010 | CHALK | Sindh |
| 150 | 01/01/2010 | CHINA CLAY | Punjab |
| 2,430 | 01/01/2010 | COPPER BLISTER | Balochistan |
| 200 | 01/01/2010 | CHROMITE | KPK |
| 22,015 | 01/01/2010 | COAL | Punjab |
| 445 | 01/01/2010 | DOLOMITE | Punjab |
| 0 | 01/01/2010 | FELD SPAR | KPK |
| 15,596 | 01/01/2010 | FIRE CLAY | Punjab |
| 46 | 01/01/2010 | FLUORITE | FATA |
| 41,725 | 01/01/2010 | GYPSUM | Punjab |
| 0 | 01/01/2010 | GRANITE | Sindh |
| 3,635 | 01/01/2010 | GRAVEL | Sindh |
| 2,335 | 01/01/2010 | IRON ORE | Punjab |
| 3,705 | 01/01/2010 | LATERITE | Punjab |
| 1,395,386 | 01/01/2010 | LIME STONE | Punjab |
| 0 | 01/01/2010 | MANGANESE | FATA |
| 0 | 01/01/2010 | MAGNISITE | KPK |
| 150 | 01/01/2010 | MARBLE | Punjab |

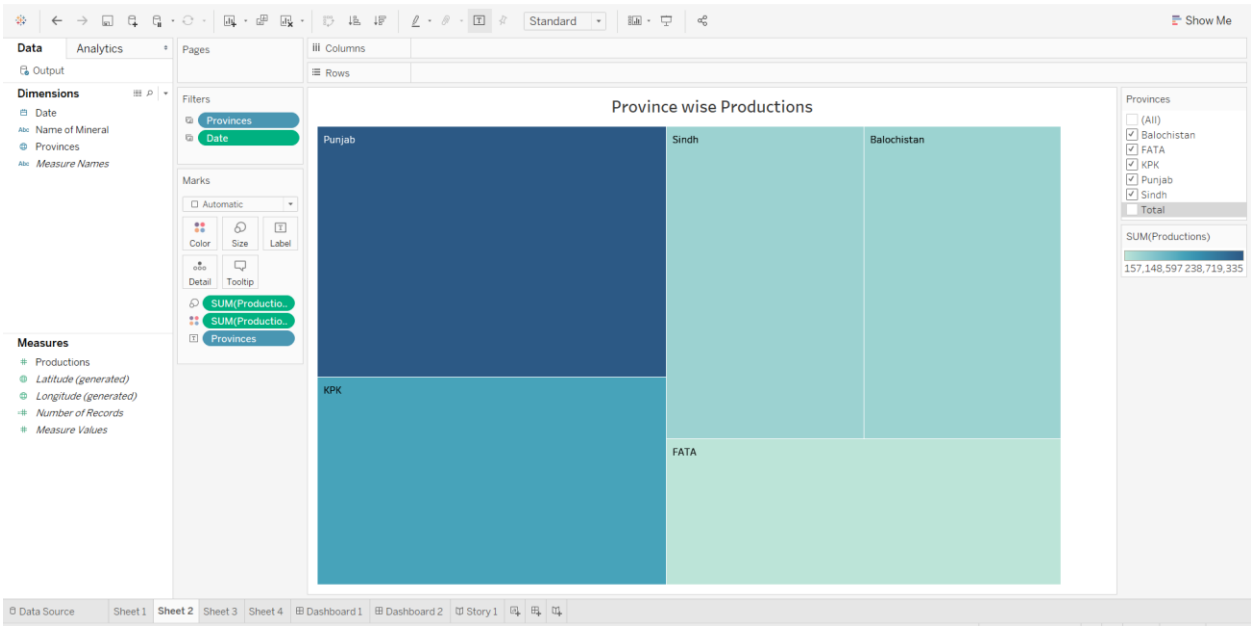
- Finally with the help of output node the cleaned and processed data was converted into csv for analysis.

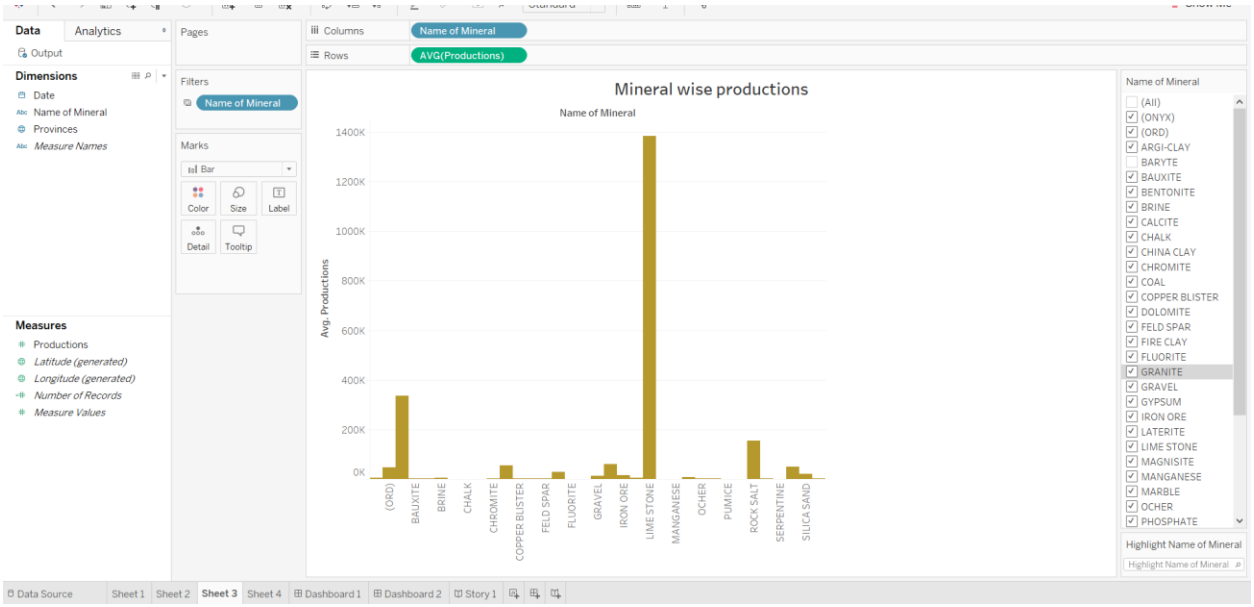
# ANALYSIS

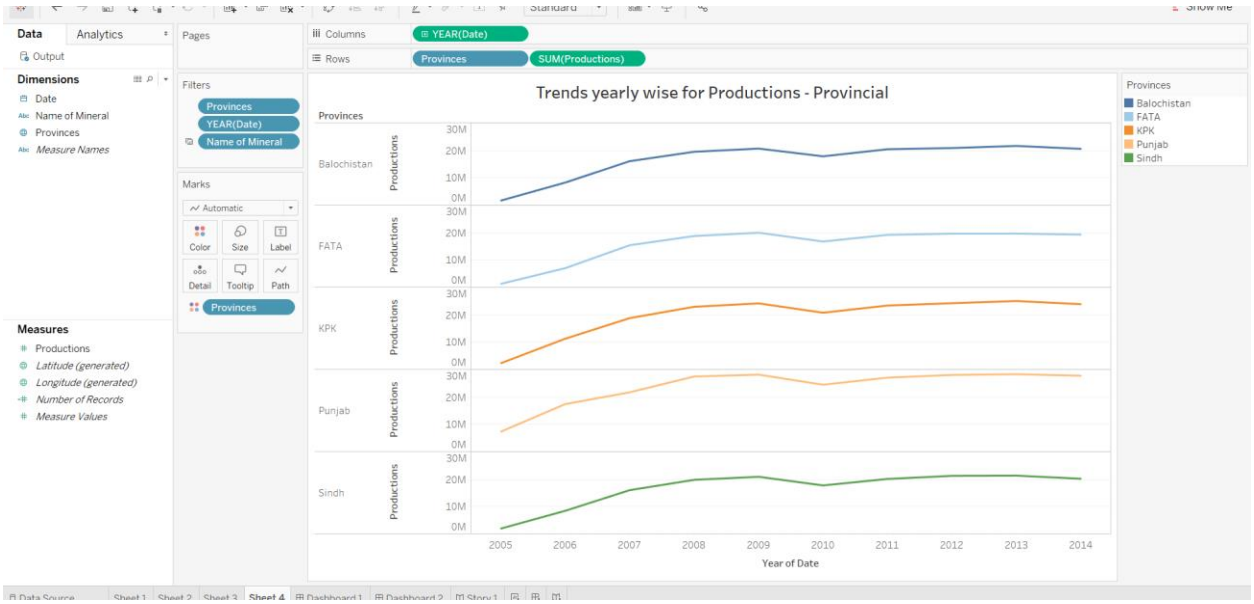
Analysis of the data was done using tableau from which I found interesting results.

Here are some snapshots from the tableau desktop:



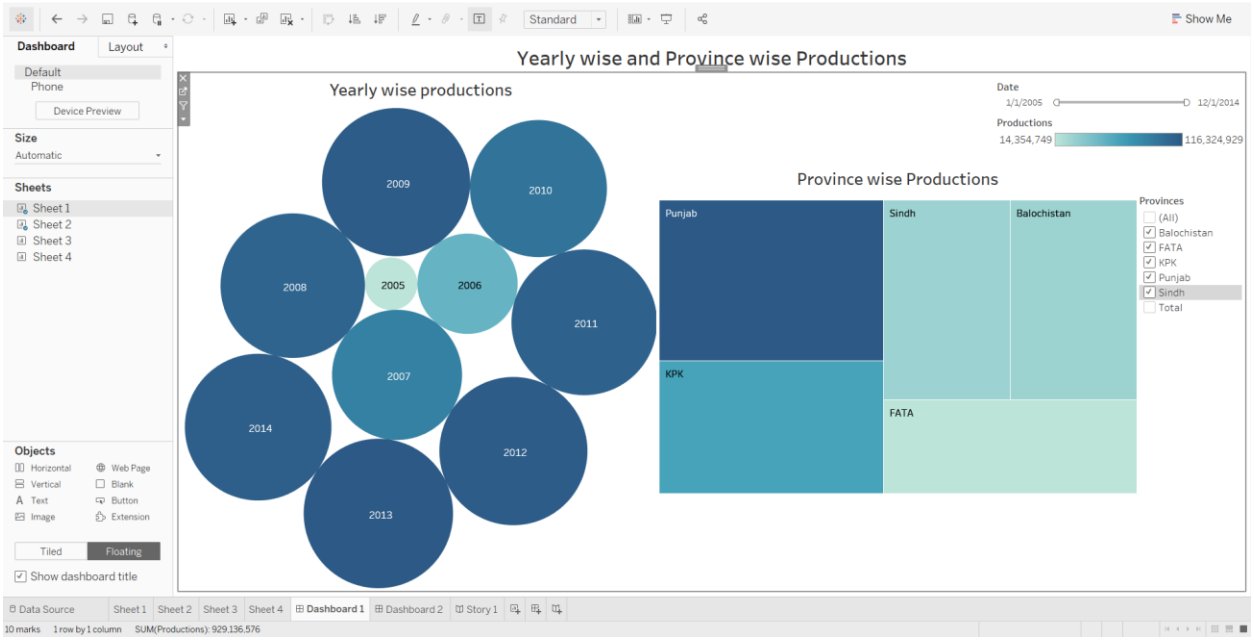





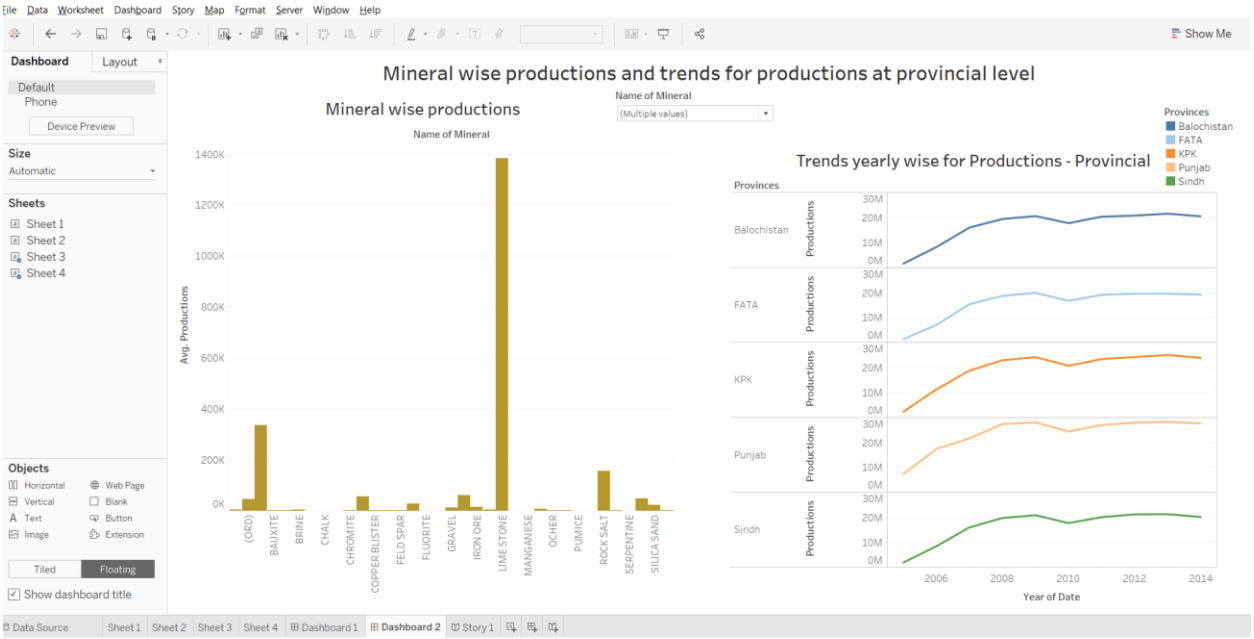




NOTE : ALL FILTERS APPLIED AT DASHBOARD ARE DYNAMIC FOR ALL THE SHEETS . ALL GRAPHS ARE CHANGING SYNCHRONOUSLY WITH THE FILTER APPLIED.

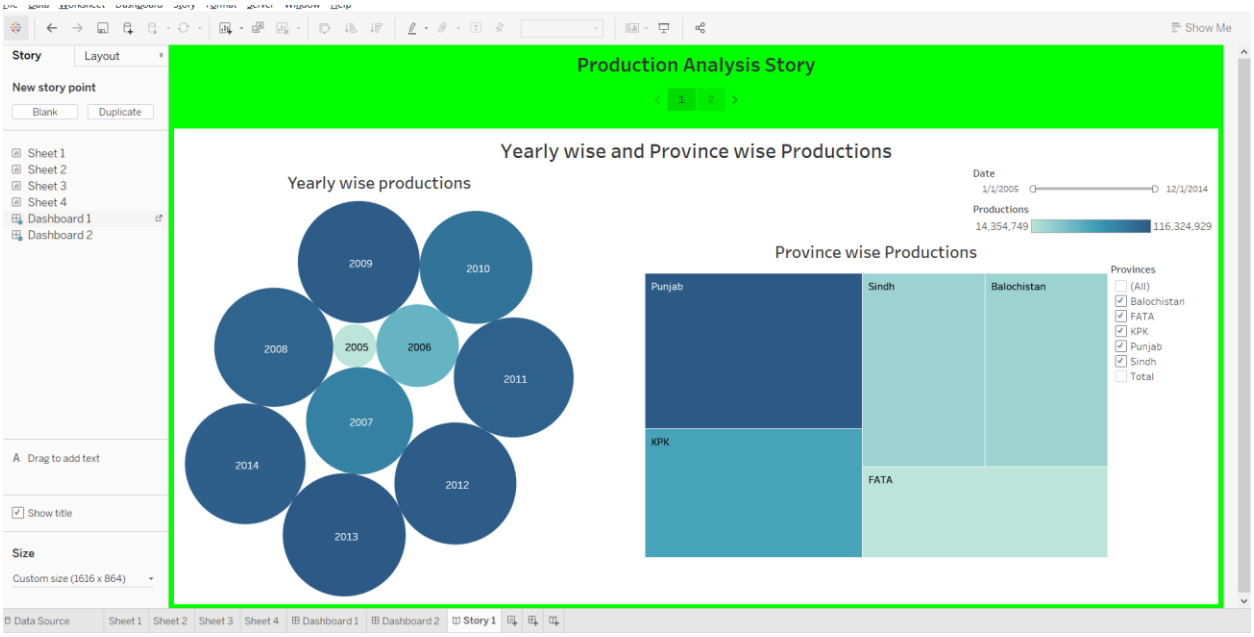
# Dashboard 1 :

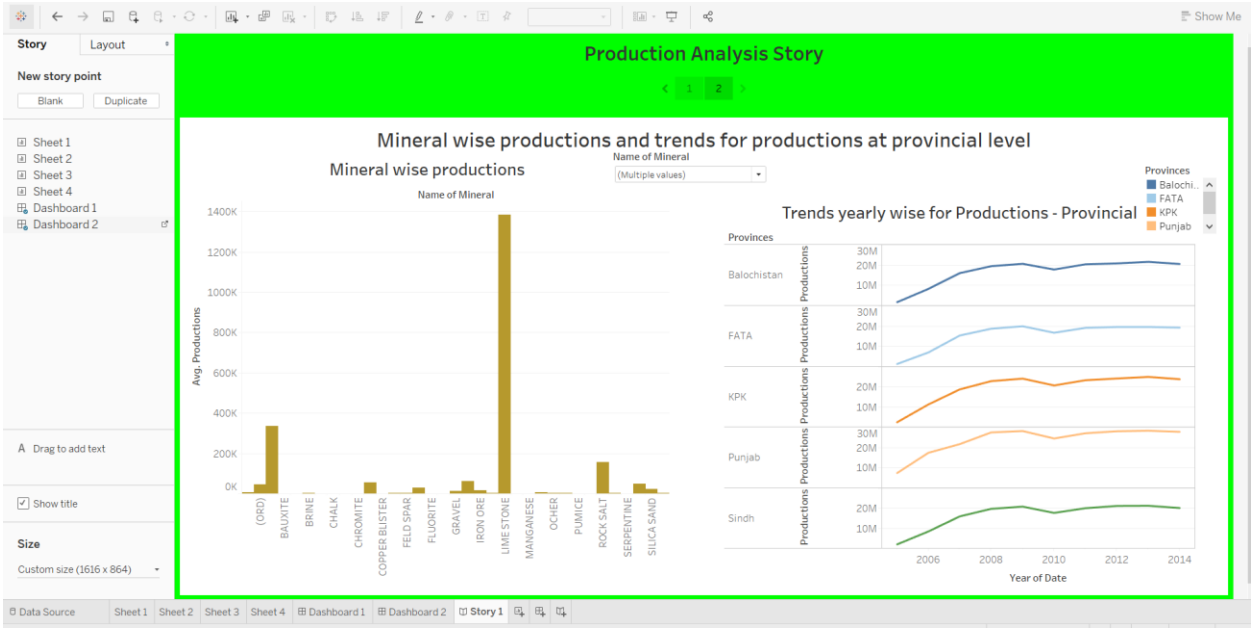



# Dashboard 2 :



# STORY :


Production Analysis Story
Yearly wise and Province wise Productions
Yearly wise productions
2009
2010
2008
2005
2006
2011
2007
2014
2012
2013
Date
1/1/2005
12/1/2014
Productions
14,354,749
116,324,929
Province wise Productions
Punjab
Sindh
Balochistan
KPK
FATA
Provinces
(All)
Balochistan
FATA
KPK
Punjab
Sindh
Total
Story
Layout
New story point
Blank
Duplicate
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Dashboard 1
Dashboard 2
Drag to add text
Show title
Size
Custom size (1616 x 864)
Show Me
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Dashboard 1
Dashboard 2
Story 1


Production Analysis Story
Mineral wise productions and trends for productions at provincial level
Mineral wise productions
Name of Mineral
(Multiple values)
Name of Mineral
Avg. Productions
1400K
1200K
1000K
800K
600K
400K
200K
0K
(ORD)
BAUXITE
BRINE
CHALK
CHROMITE
COPPER BLISTER
FELD SPAR
FLUORITE
GRAVEL
IRON ORE
LIME STONE
MANGANESE
OCHER
PUMICE
ROCK SALT
SERPENTINE
SILICA SAND
Trends yearly wise for Productions - Provincial
Provinces
Balochi..
FATA
KPK
Punjab
Balochistan
FATA
KPK
Punjab
Sindh
Productions
30M
20M
10M
2006
2008
2010
2012
2014
Year of Date
Story
Layout
New story point
Blank
Duplicate
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Dashboard 1
Dashboard 2
Drag to add text
Show title
Size
Custom size (1616 x 864)
Show Me
Data Source
Sheet 1
Sheet 2
Sheet 3
Sheet 4
Dashboard 1
Dashboard 2
Story 1

# RESULTS :

This was the analysis for the Energy and mining statistics, we found interesting results from bubble chart and tree map that most of the productions took place in 2013 and the majority belonged to Punjab while the least took place in 2005 and minority were in FATA province. Considering mineral productions limestone won the game of having the highest productions and from trend analysis we found different trends like the highest peak of production in balochistan and FATA was in 2009 and others had their highest peaks at 2013.