

Tableau

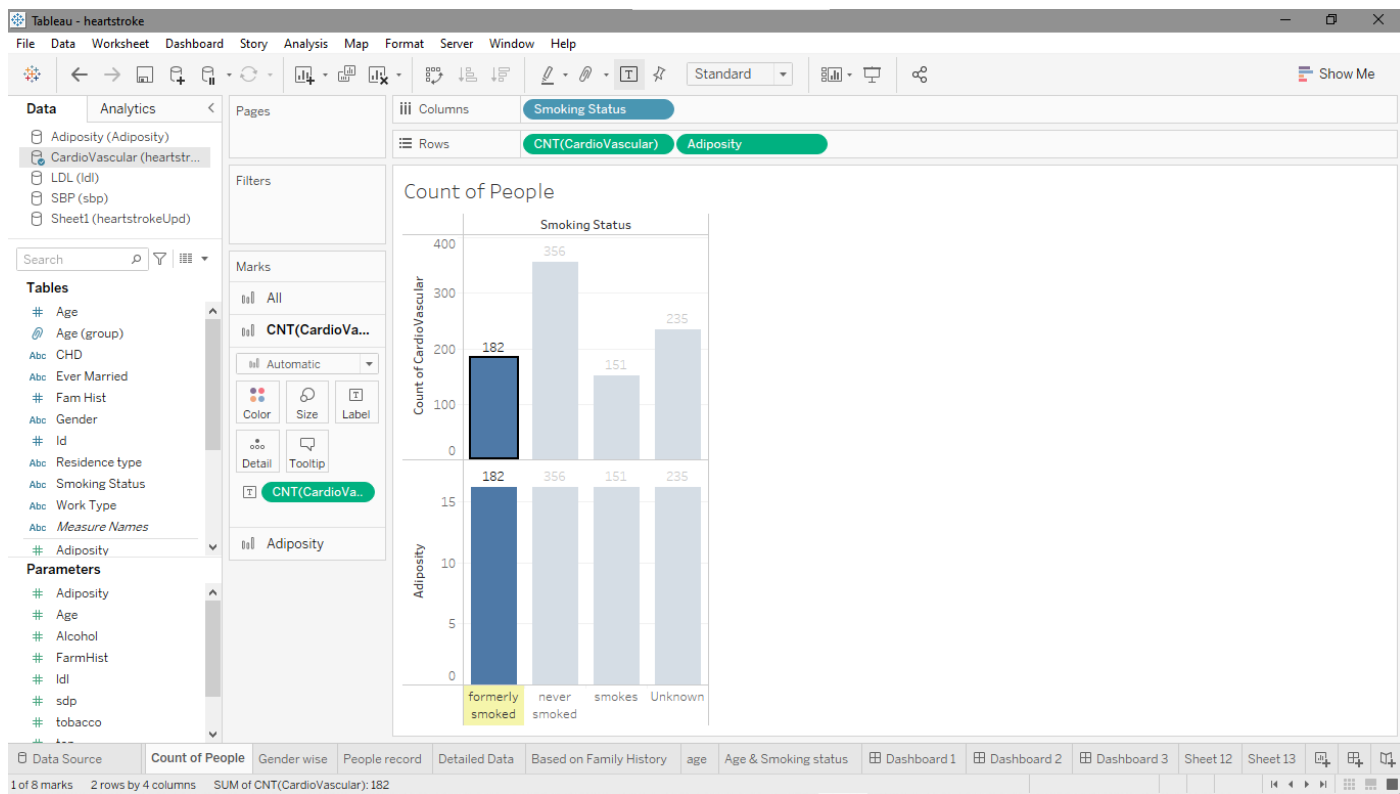
by Vishal Kumar Singh
MCA-ss (2019272041)

Heatstroke/ Heart disease Prediction and Analysis Using Tableau

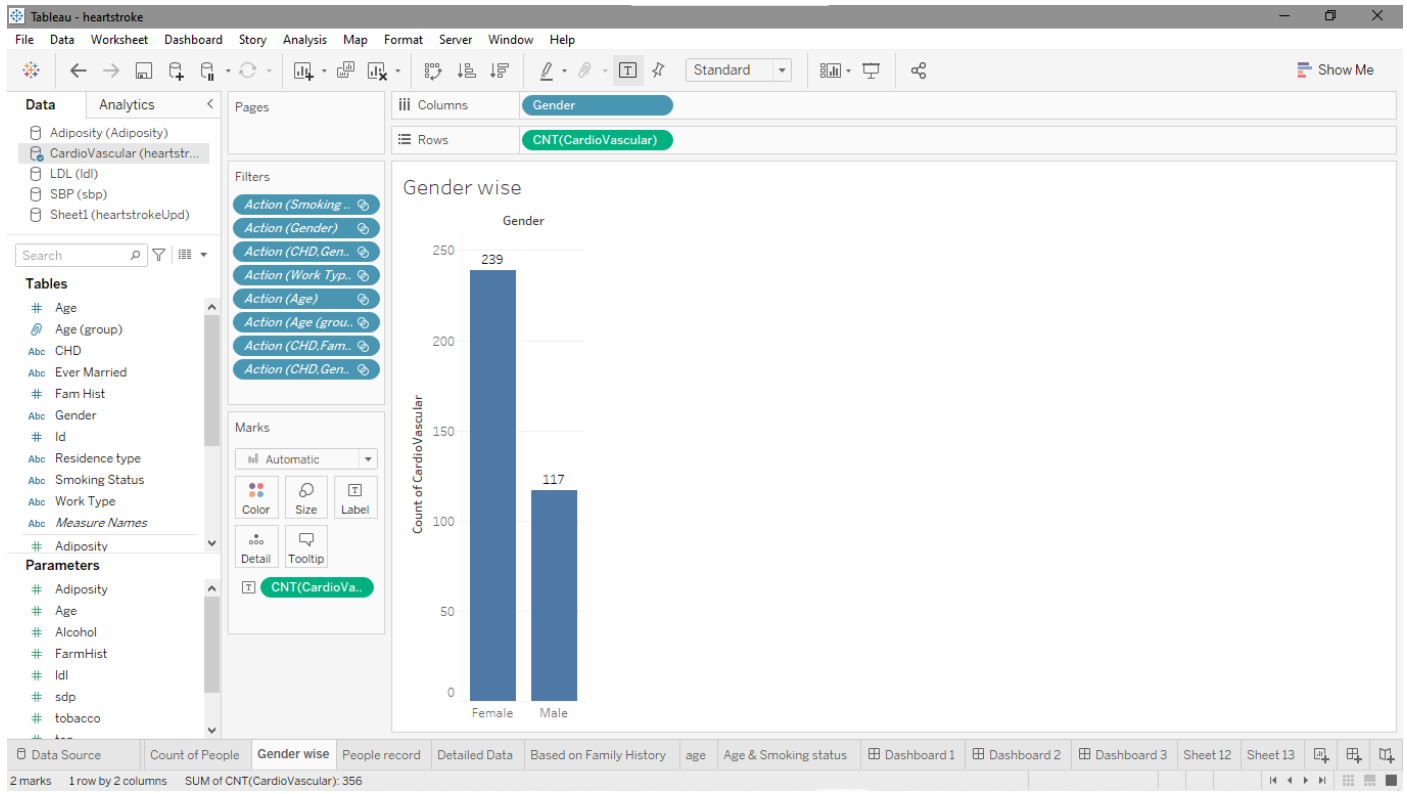
Analysis of the dataset

The screenshot shows the Tableau 'heartstroke' workbook. The 'Connections' pane on the left lists 'heartstrokeUpd' as a Microsoft Excel source. The 'Sheets' pane shows 'Sheet1' and 'New Union'. The main view is 'Sheet1 (heartstrokeUpd)', which is currently empty, displaying a 'Need more data?' message. Below this, a table of fields is visible, including 'Adiposity', 'tobacco(in Kg)', 'Alcohol', 'Typea', 'Age', 'Fam Hist', and 'CHD'. The bottom status bar indicates '9 fields 1 rows'.

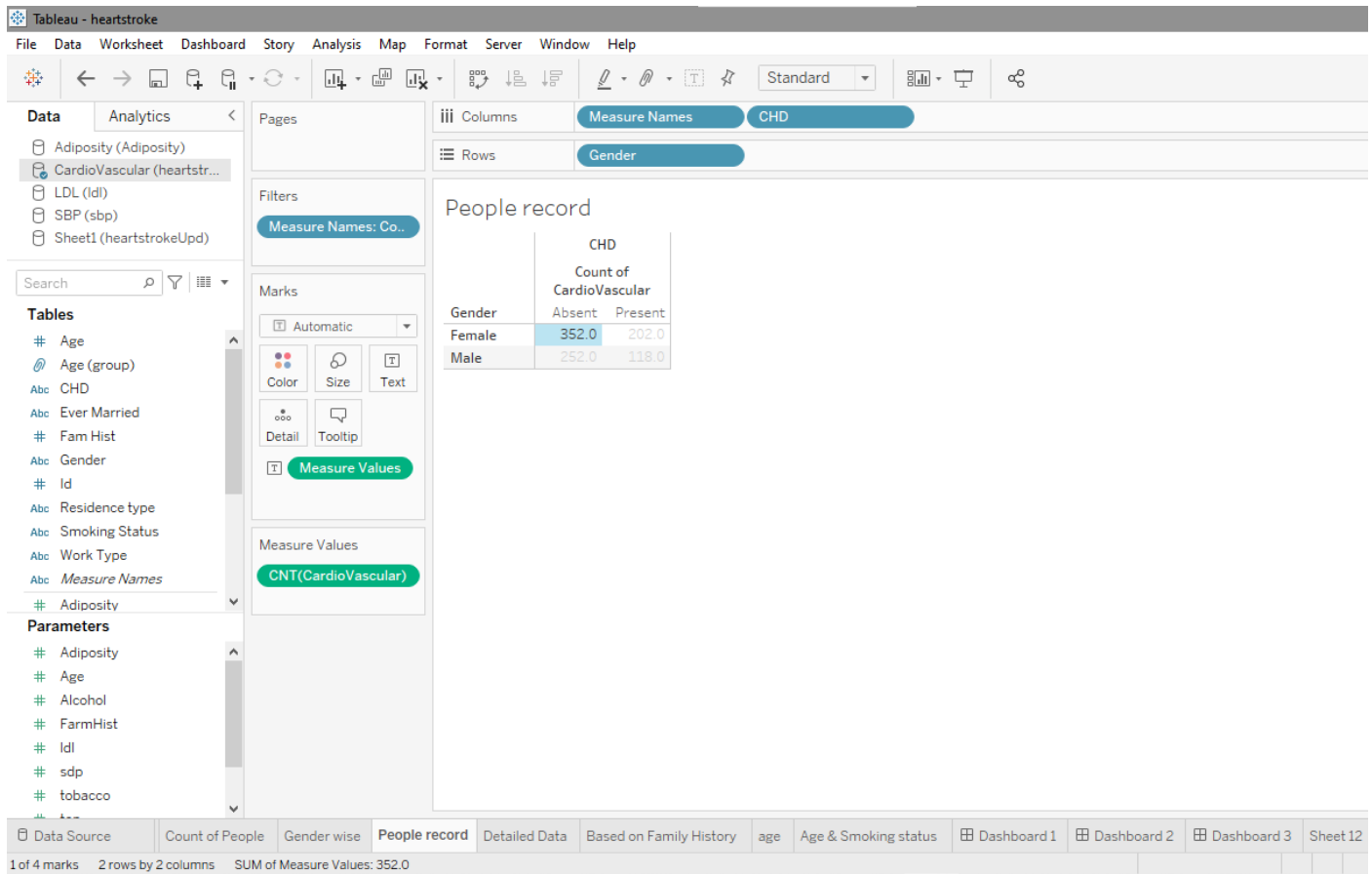
- Count of people based on cardiovascular and adiposity



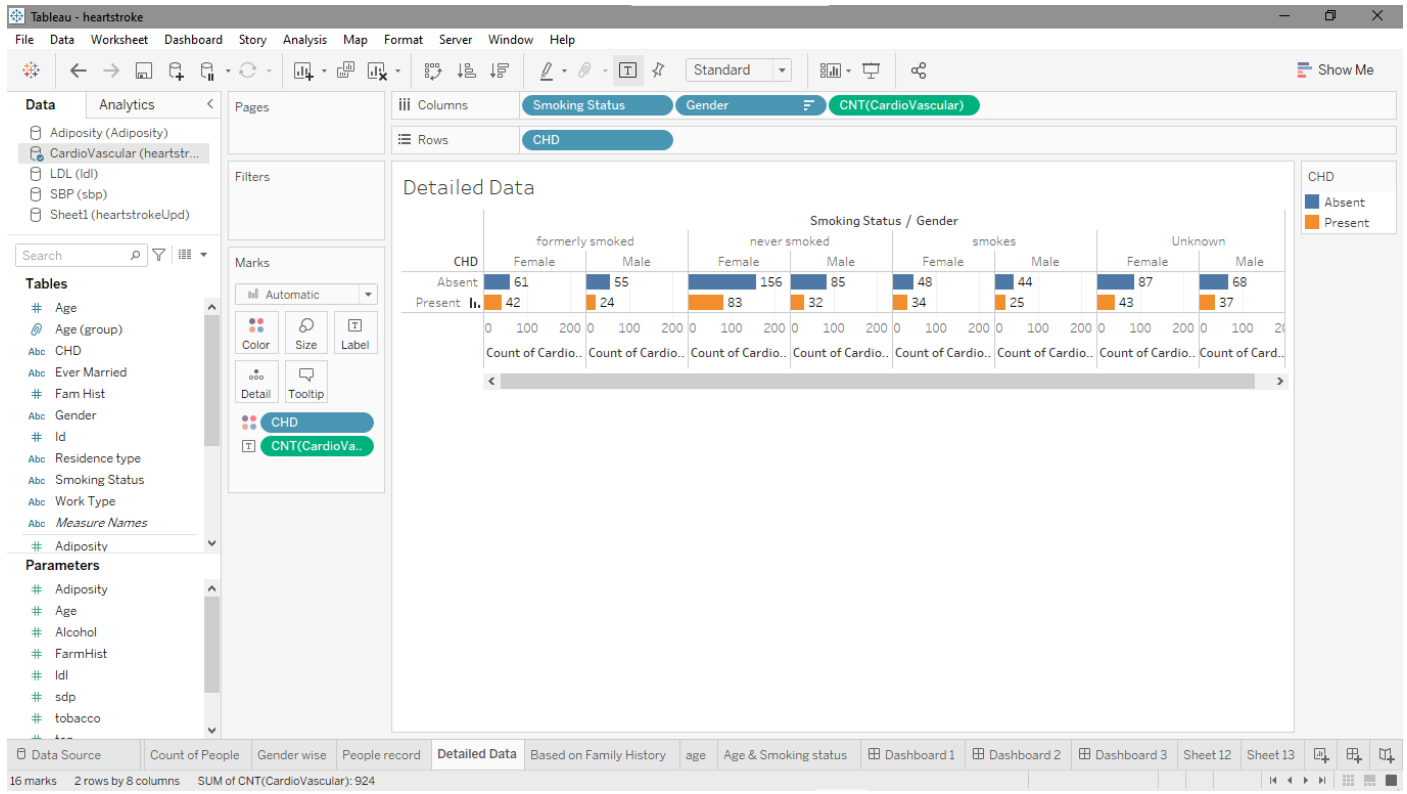
- Grouping the count gender wise



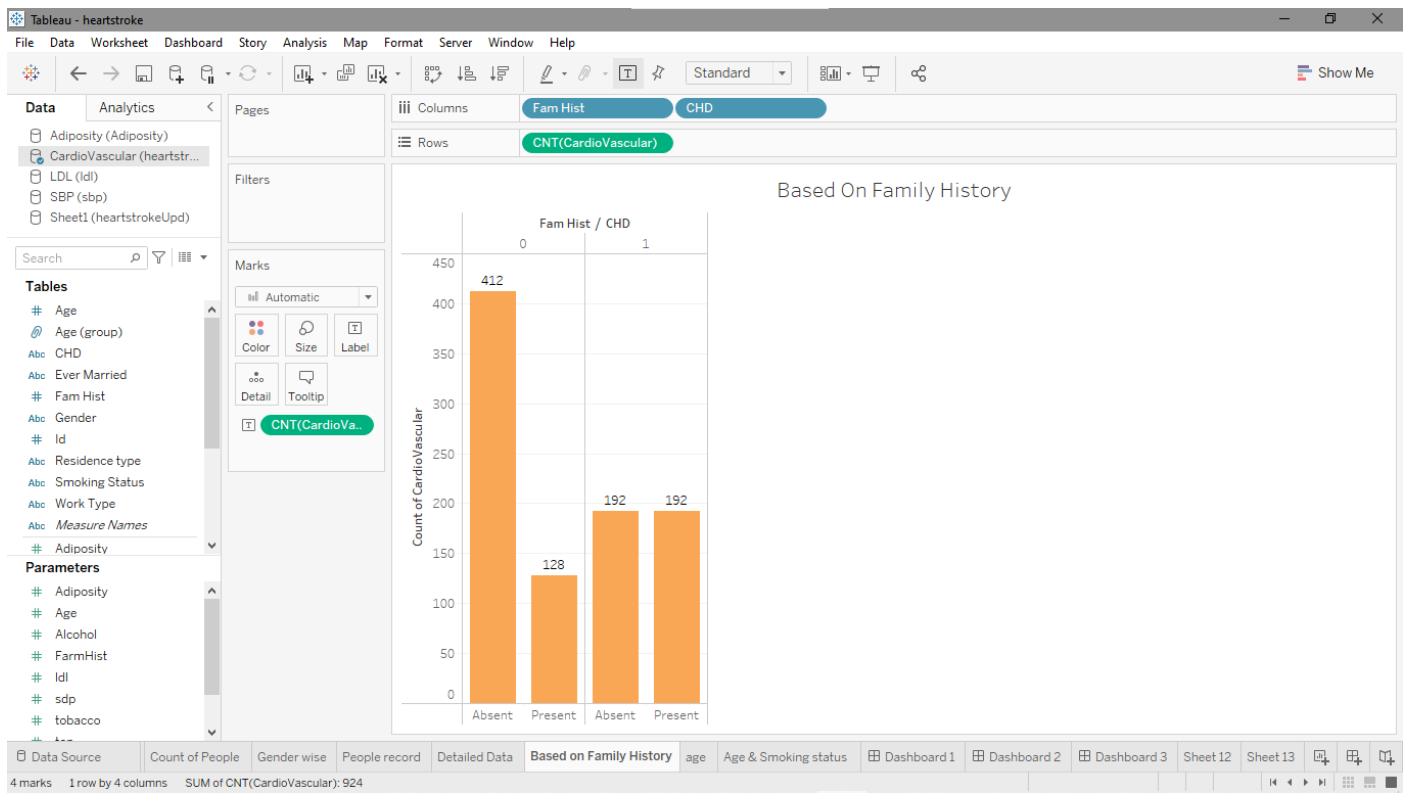
- Calculation of the people records gender wise



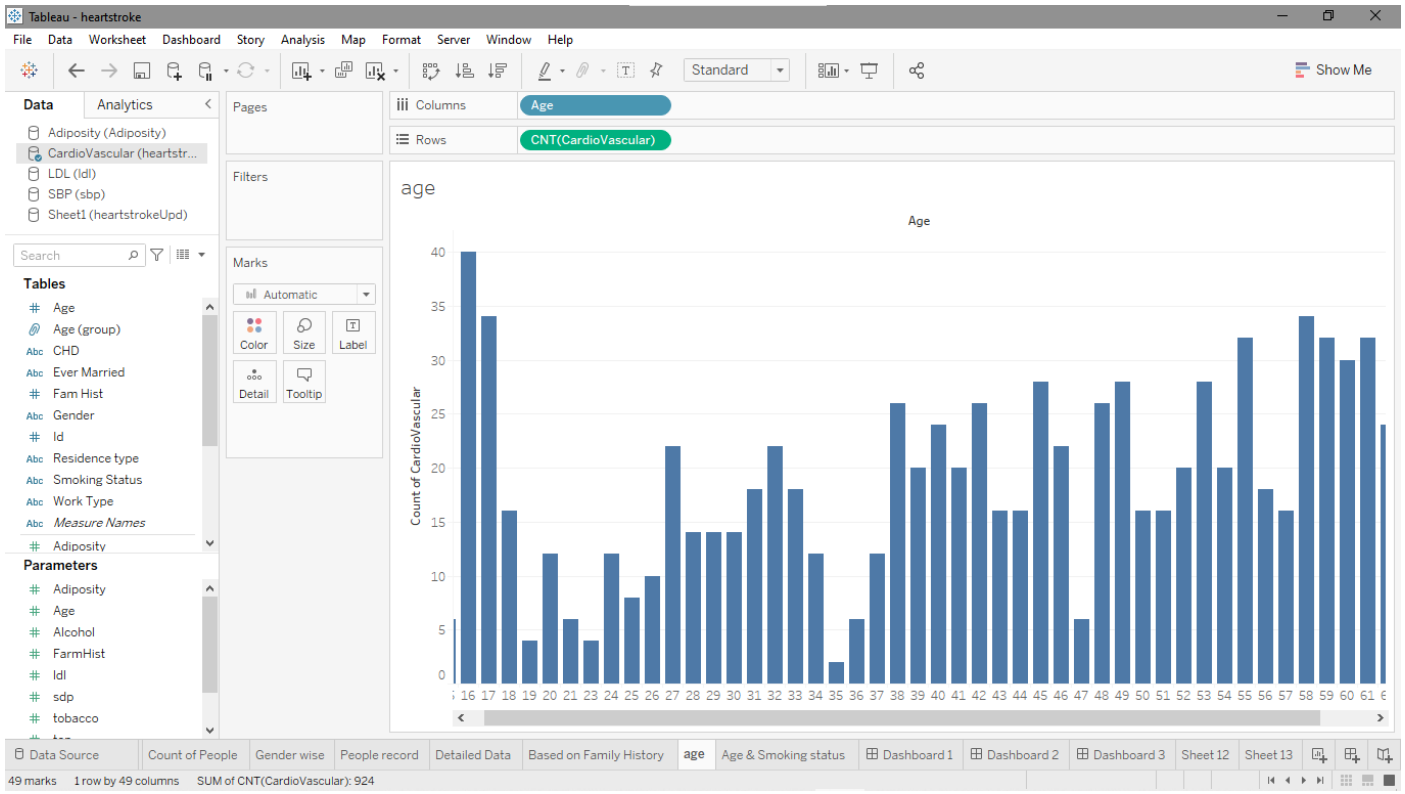
- Detailed data for each categories



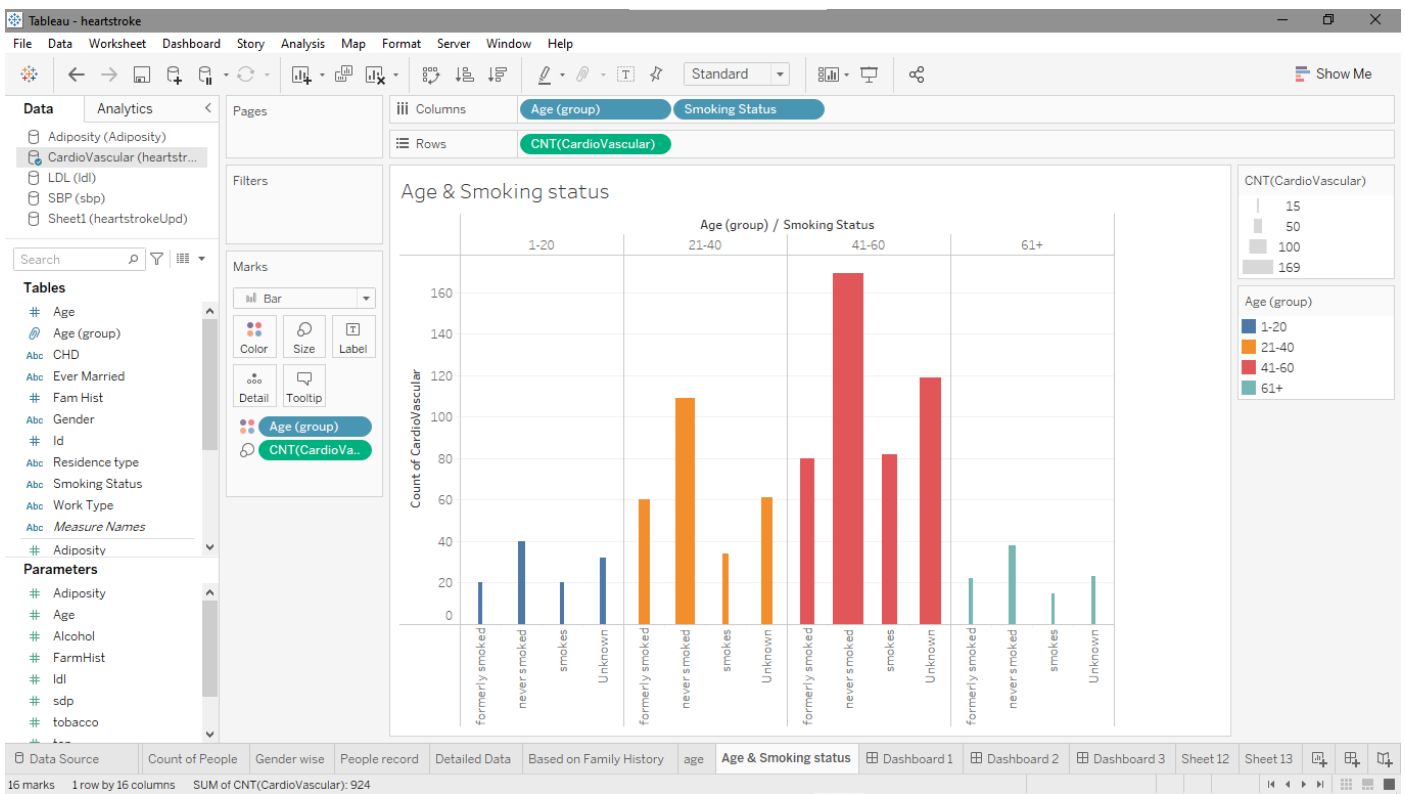
- Analysis based on family history



- Merging and grouping based on age



- Merging and grouping based on age and smoking status



Predictive modelling(Using Logistic Regression)

- Model for % of presence

```
SCRIPT_STR(  
"  
import pandas as pd  
import numpy as np  
from ast import literal_eval  
from sklearn.linear_model import LogisticRegression  
from sklearn.preprocessing import StandardScaler  
  
d={  
'ad':literal_eval(_arg1[0]),  
'ag':literal_eval(_arg2[0]),  
'al':literal_eval(_arg3[0]),  
'fam':literal_eval(_arg4[0]),  
'ldl':literal_eval(_arg5[0]),  
'sbp':literal_eval(_arg6[0]),  
'tob':literal_eval(_arg7[0]),  
'typea':literal_eval(_arg8[0]),  
'chd':literal_eval(_arg9[0])  
}  
  
df=pd.DataFrame(d)  
X=df.iloc[:,8]  
y=df.iloc[:,8]  
  
sc=StandardScaler()  
X_t=sc.fit_transform(X)  
  
lr=LogisticRegression()  
lr.fit(X_t,y)  
  
inp_list=[_arg10[0],_arg11[0],_arg12[0],_arg13[0],  
_arg14[0],_arg15[0],_arg16[0],_arg17[0]]  
inp=np.array(inp_list).reshape(1,-1)  
inp=sc.transform(inp)  
pred=lr.predict(inp)  
prob=lr.predict_proba(inp)  
  
return str(pred[0])  
",  
ATTR([Adiposity]),ATTR([Age]),ATTR([Alcohol]),ATTR([Fam Hist]),  
ATTR([ldl(mmol/L)]),ATTR([sbp(mm Hg)]),ATTR([tobacco(in Kg)]),ATTR([Typea]),  
ATTR([CHD]),  
[Parameters].[Adiposity],[Parameters].[Age],[Parameters].[Alcohol],  
[FarmHist],[ldl],[sdp],[tobacco],[Parameters].[Typea]  
)
```

- Model for % of absence

```

FLOAT(SCRIPT_STR(
"
import pandas as pd
import numpy as np
from ast import literal_eval
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler

d={
'ad':literal_eval(_arg1[0]),
'ag':literal_eval(_arg2[0]),
'al':literal_eval(_arg3[0]),
'fam':literal_eval(_arg4[0]),
'ldl':literal_eval(_arg5[0]),
'sbp':literal_eval(_arg6[0]),
'tob':literal_eval(_arg7[0]),
'typea':literal_eval(_arg8[0]),
'chd':literal_eval(_arg9[0])
}

df=pd.DataFrame(d)
X=df.iloc[:, :8]
y=df.iloc[:, 8]

sc=StandardScaler()
X_t=sc.fit_transform(X)

lr=LogisticRegression()
lr.fit(X_t,y)

inp_list=[_arg10[0],_arg11[0],_arg12[0],_arg13[0],
_arg14[0],_arg15[0],_arg16[0],_arg17[0]]

inp=np.array(inp_list).reshape(1,-1)
inp=sc.transform(inp)
pred=lr.predict(inp)
prob=lr.predict_proba(inp)

return str(prob[0][1])
",

ATTR([Adiposity]),ATTR([Age]),ATTR([Alcohol]),ATTR([Fam Hist]),
ATTR([ldl(mmol/L)]),ATTR([sbp(mm Hg)]),ATTR([tobacco(in Kg)]),ATTR([Typea]),
ATTR([CHD]),
[Parameters].[Adiposity],[Parameters].[Age],[Parameters].[Alcohol],
[FarmHist],[ldl],[sdp],[tobacco],[Parameters].[Typea]
))

```

- Model for getting accuracy

```

SCRIPT_STR(
"
import pandas as pd
import numpy as np
from ast import literal_eval
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler

d={
'ad':literal_eval(_arg1[0]),
'ag':literal_eval(_arg2[0]),
'al':literal_eval(_arg3[0]),
'fam':literal_eval(_arg4[0]),
'ldl':literal_eval(_arg5[0]),
'sbp':literal_eval(_arg6[0]),
'tob':literal_eval(_arg7[0]),
'typea':literal_eval(_arg8[0]),
'chd':literal_eval(_arg9[0])
}

df=pd.DataFrame(d)
X=df.iloc[:, :8]
y=df.iloc[:, 8]

sc=StandardScaler()
X_t=sc.fit_transform(X)

lr=LogisticRegression()
lr.fit(X_t,y)
return str(lr.score(X_t,y))
",

ATTR([Adiposity]),ATTR([Age]),ATTR([Alcohol]),ATTR([Fam Hist]),
ATTR([ldl(mmol/L)]),ATTR([sbp(mm Hg)]),ATTR([tobacco(in Kg)]),ATTR([Typea]),
ATTR([CHD]),
[Parameters].[Adiposity],[Parameters].[Age],[Parameters].[Alcohol],
[FarmHist],[ldl],[sdp],[tobacco],[Parameters].[Typea]
)

```


- Model for getting presence or absence

```

SCRIPT_STR(
"
import pandas as pd
import numpy as np
from ast import literal_eval
from sklearn.linear_model import LogisticRegression
from sklearn.preprocessing import StandardScaler

d={
'ad':literal_eval(_arg1[0]),
'ag':literal_eval(_arg2[0]),
'al':literal_eval(_arg3[0]),
'fam':literal_eval(_arg4[0]),
'ldl':literal_eval(_arg5[0]),
'sbp':literal_eval(_arg6[0]),
'tob':literal_eval(_arg7[0]),
'typea':literal_eval(_arg8[0]),
'chd':literal_eval(_arg9[0])
}

df=pd.DataFrame(d)
X=df.iloc[:,8]
y=df.iloc[:,8]

sc=StandardScaler()
X_t=sc.fit_transform(X)

lr=LogisticRegression()
lr.fit(X_t,y)

inp_list=[_arg10[0],_arg11[0],_arg12[0],_arg13[0],
_arg14[0],_arg15[0],_arg16[0],_arg17[0]]

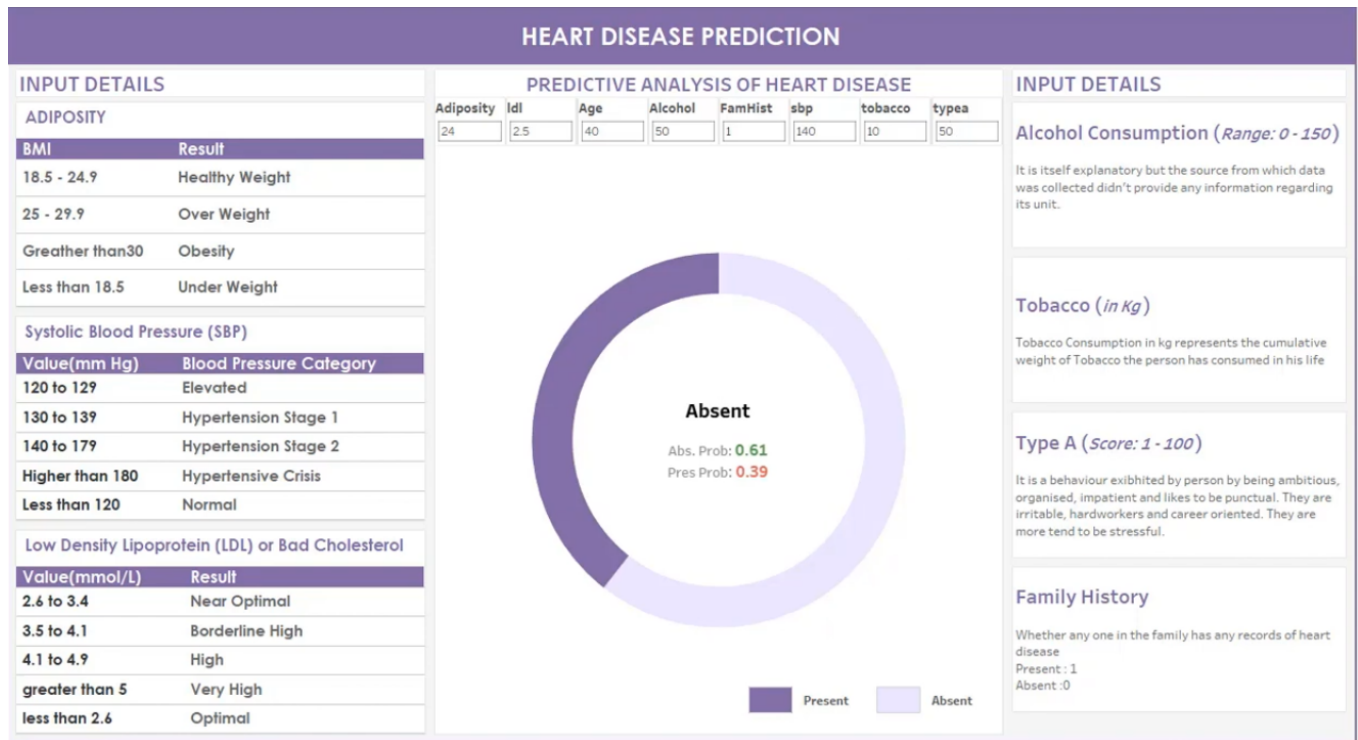
inp=np.array(inp_list).reshape(1,-1)
inp=sc.transform(inp)
pred=lr.predict(inp)
prob=lr.predict_proba(inp)

return str(pred[0])
",

ATTR([Adiposity]),ATTR([Age]),ATTR([Alcohol]),ATTR([Fam Hist]),
ATTR([ldl(mmol/L)]),ATTR([sbp(mm Hg)]),ATTR([tobacco(in Kg)]),ATTR([Typea]),
ATTR([CHD]),
[Parameters].[Adiposity],[Parameters].[Age],[Parameters].[Alcohol],
[FamHist],[ldl],[sdp],[tobacco],[Parameters].[Typea]
)

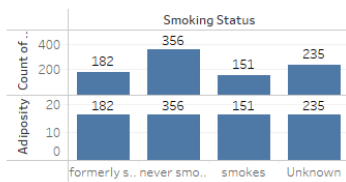
```

Designed Model in work

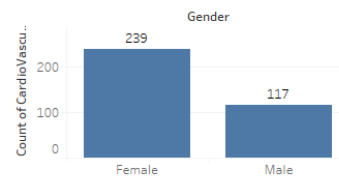


• Dashboard 1

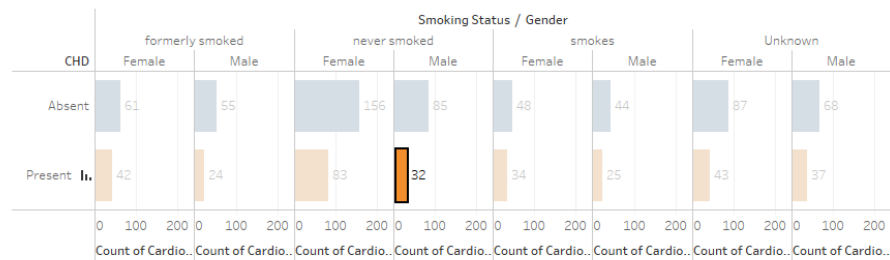
Count of People



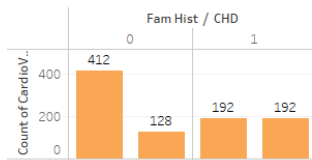
Gender wise



Detailed Data



Based On Family History



• Dashboard 2

Age & Smoking status

