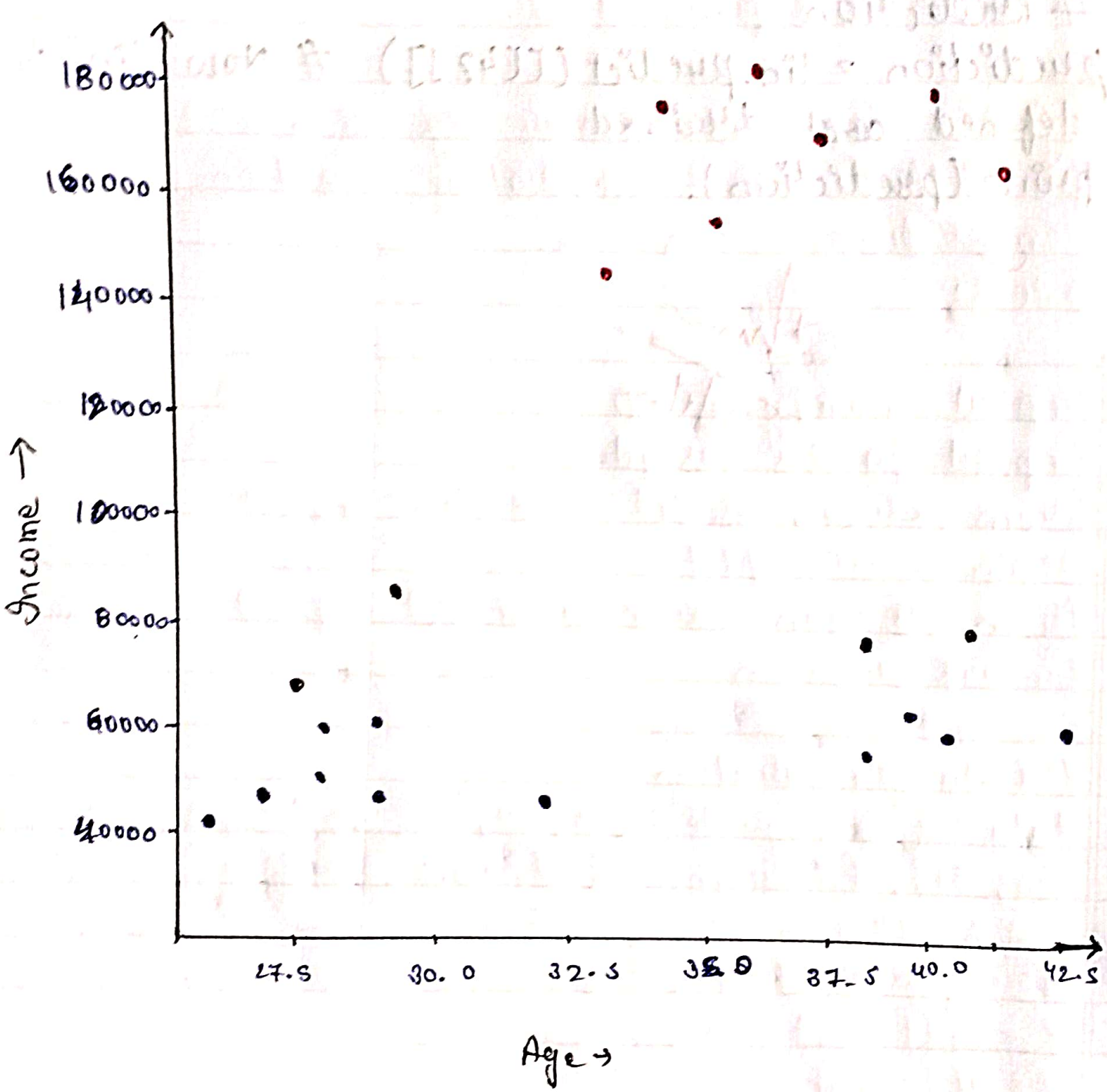

180000
160000
140000
120000
100000
80000
60000
40000
Income →
27.5
30.0
32.5
35.0
37.5
40.0
42.5
Age →

POORNIMA

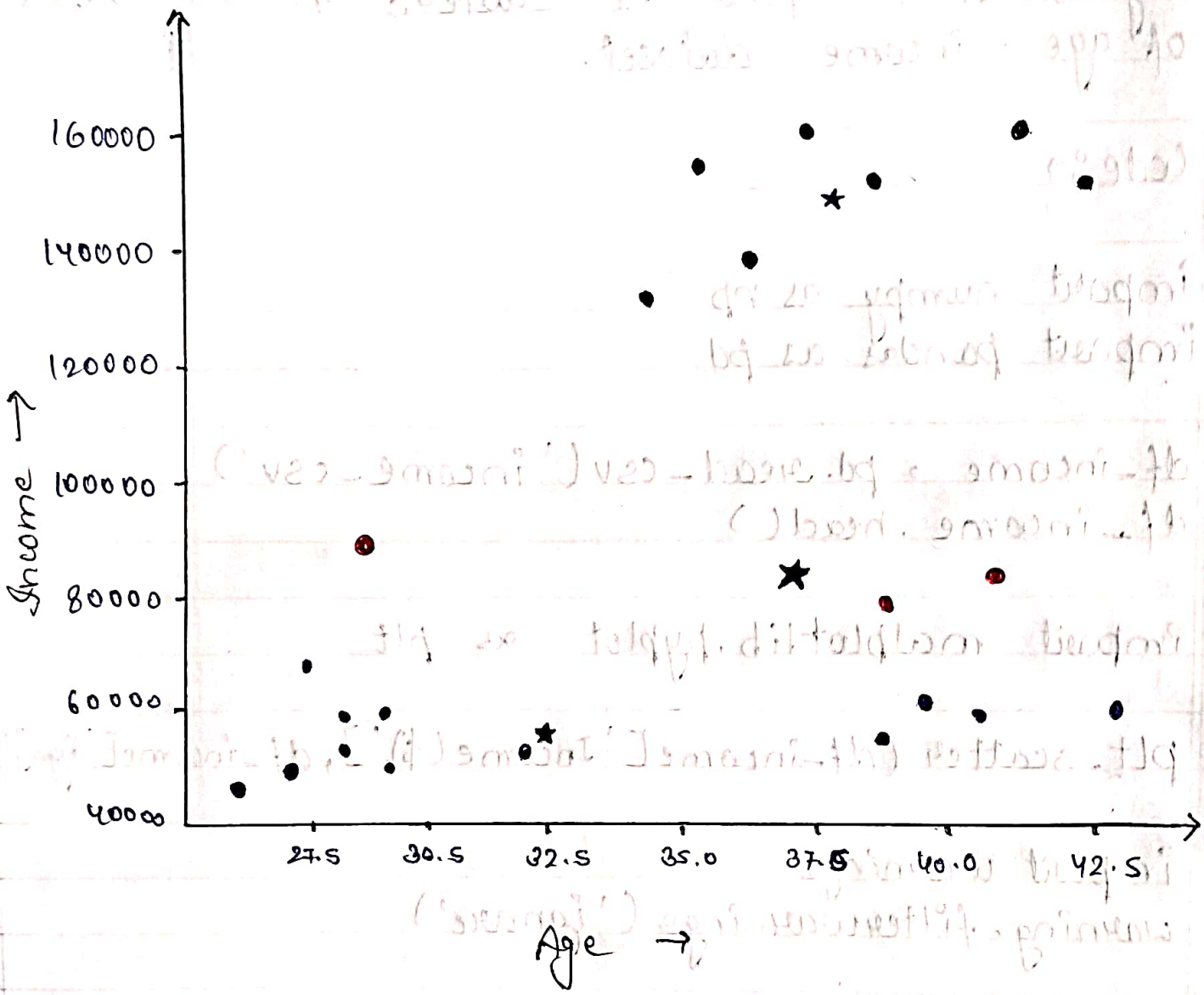
Experiment - 3

- Objective :→ Implement K-means clustering Algorithm for find the clusters on the basis of age - income dataset.

- Code :→

```
import numpy as np
import pandas as pd

df_income = pd.read_csv('income_csv')
df_income.head()

import matplotlib.pyplot as plt

plt.scatter(df_income['Income($)'], df_income['Age'])

import warnings
warning.filterwarnings('ignore')

from sklearn.cluster import KMeans
Km = KMeans(n_clusters = 3)

y_predicted_clusters = km.fit_predict(df_income[['Age',
'Income($)']])

y_predicted_clusters.
```

Page No.....


160000
140000
120000
100000
80000
60000
40000
Income →
27.5
30.5
32.5
35.0
37.5
40.0
42.5
Age →

POORNIMA

```
df.income['cluster'] = y_predicted_clusters

df_income.head()

plt.scatter(df_income['Age'], df_income['Income($)']
,c=df_income['cluster'], cmap='rainbow')

y_predicted_clusters = km.fit_predict(df_income[[
    'Age','Income($)']])

df_income['Cluster'] = y_predicted_cluster

plt.scanner(df_income['Age'], df_income['Income($)']
c=df_income['cluster'], cmap='rainbow').
plt.xlabel('Age')
plt.ylabel('Income in $')

cluster_centers = km.cluster_centers_

plt.scatter(df_income['Age'], df_income('Income($)']
c=df_income['cluster'], cmap='rainbow')
plt.scatter(cluster_centers[:,0], cluster_centers[:,1],
color='purple', marker='*')
plt.xlabel('Age')
plt.ylabel('Income in $')

k_range = range(1, 11)
sse = []
```

Page No.....

POORNIMA

```
for k in k_range:
    km = KMeans(n_clusters=k)
    km.fit_predict(df_income[['Age', 'Income($)']])
    sse.append(km.inertia_)

plt.xlabel('k')
plt.ylabel('SSE')
plt.plot(k_range, sse).
```