

# CLUSTERING AND PCA ASSIGNMENT

## SUBMISSION

Group Name: rudra

1. Member name    rudra pratap sen

## <Abstract>

- **Introduction** – the project is about a NGO named HELP INTERNATIONAL it's help countries to fight poverty.
- **Problem statement** - .the CEO of the company wants to use the money strategically and effectively which he got in recent funding . And want to donate in those countries which are direst need of it
- **Procedures** – we have got a data of countries and has some factor like gdpp , child mort , health , income,inflation etc. we have to choose socio economic factor which describe country socio economic condition
- **Results** – we have got 5 countries name which has very low socio economic factor

## <Analysis>

### Analysis of principal components

After standard scaling of data , we do pca projection on data and found that 90% data is explained by 3 principal components . And also cumulative explained variance graph is also showing the same.

We choose the first 3 principal components and made a scatter plot which feature is mostly defined by these principal Components.

## <Analysis>

### Analysis of k-means clustering

At first we have done silhouette analysis and checked sum of squared distance graphs , from both the graphs

We conclude that we should choose 5 cluster for data . As we choose 4 main feature for socio economic  
i.e child\_mort,income,inflation,gdpp.

As cluster 2 and cluster 4 countries have low values among them sowe can conclude both cluster countries in final list  
of countries

## <Analysis>

### Analysis of the hierarchical clustering

We do two way of hierarchical clustering

1. Single
2. Complete

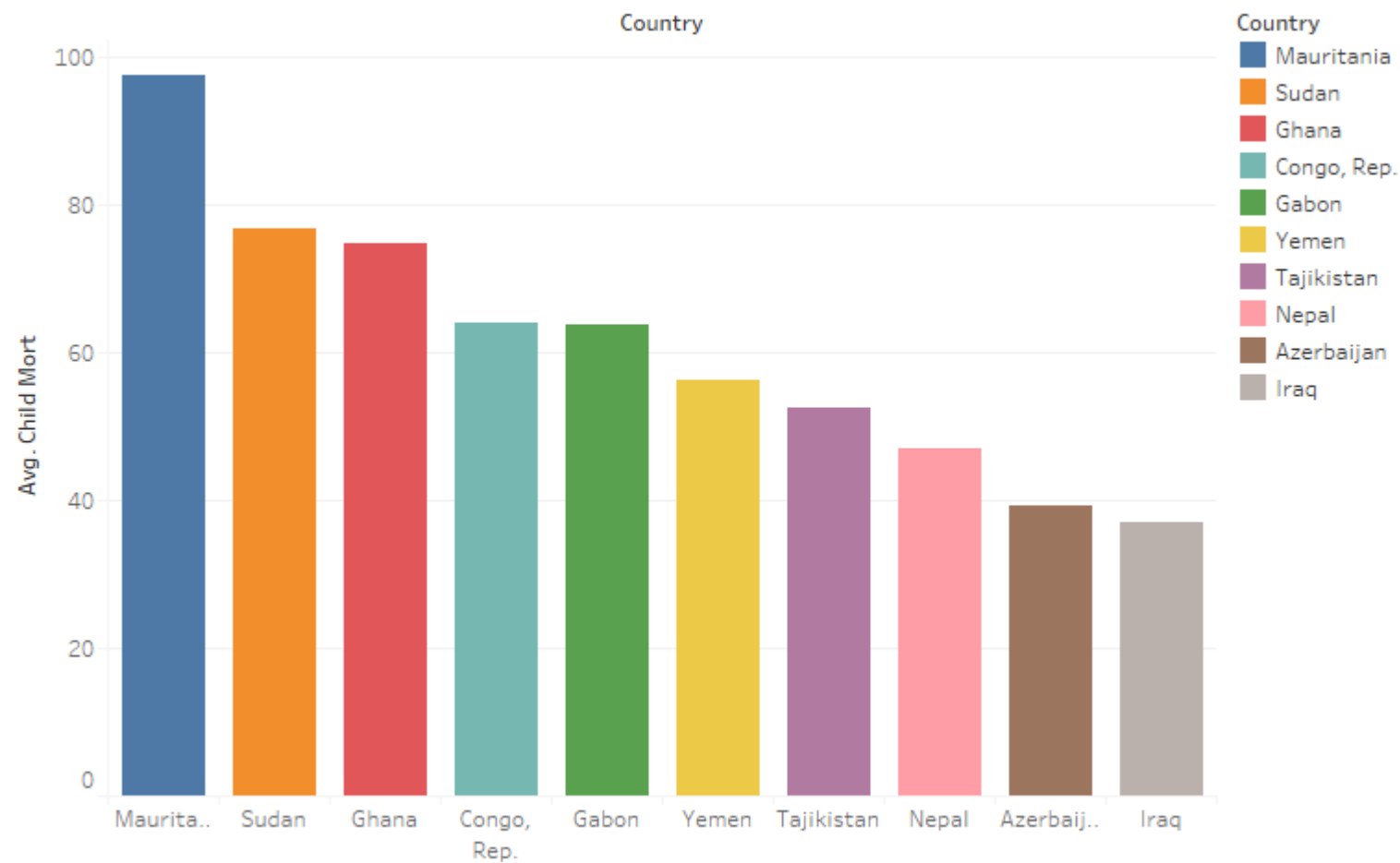
Single is not giving good result so we choose 5 cluster with the help of complete method

We choose cluster no 4 from hierarchical cluster and merge in final list of countries

# <Results>

Sheet 1

Plot 1

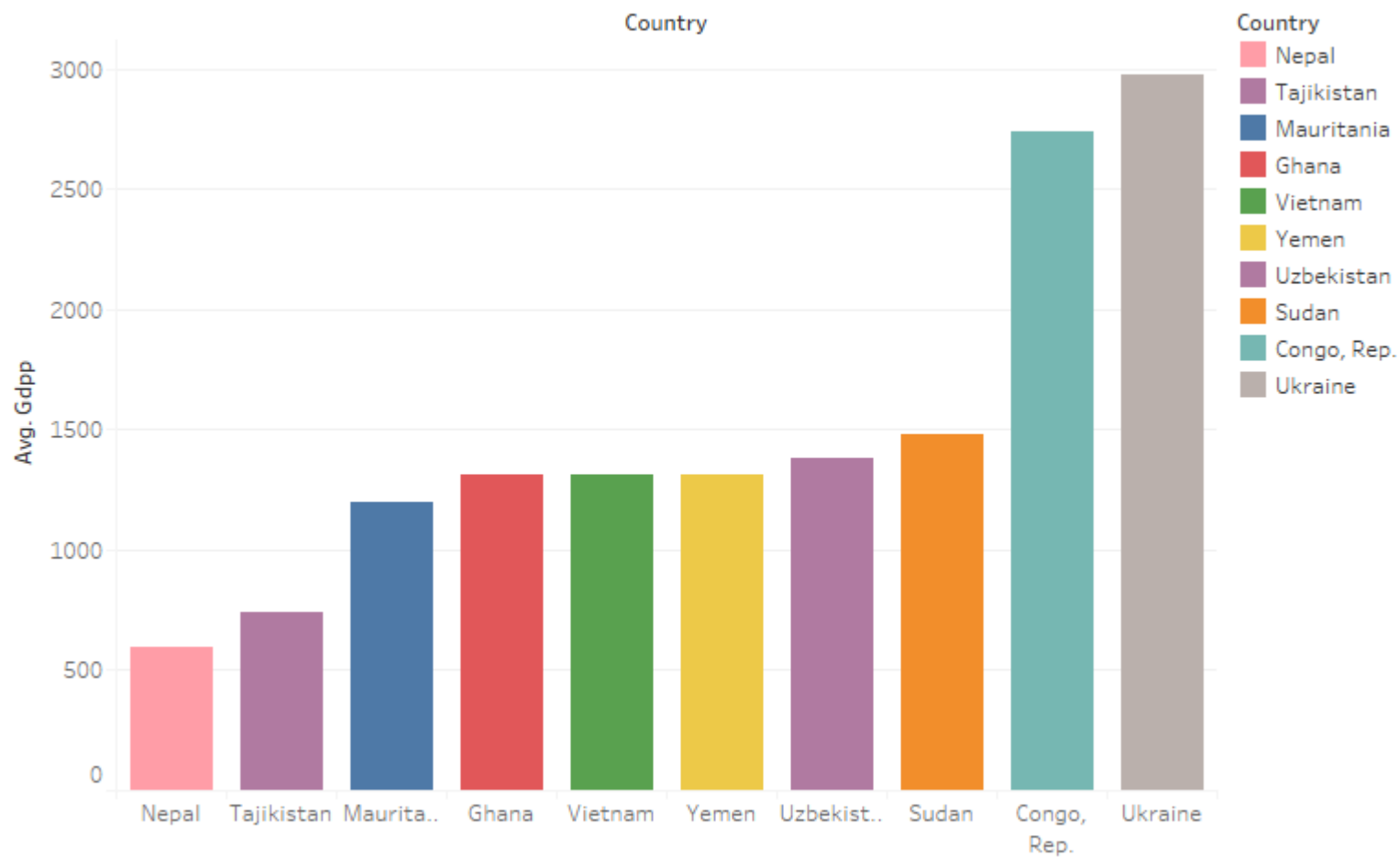


Average of Child Mort for each Country. Color shows details about Country. The view is filtered on Country, which keeps 10 of 25 members.

# <Results>

Plot 2

Sheet 2

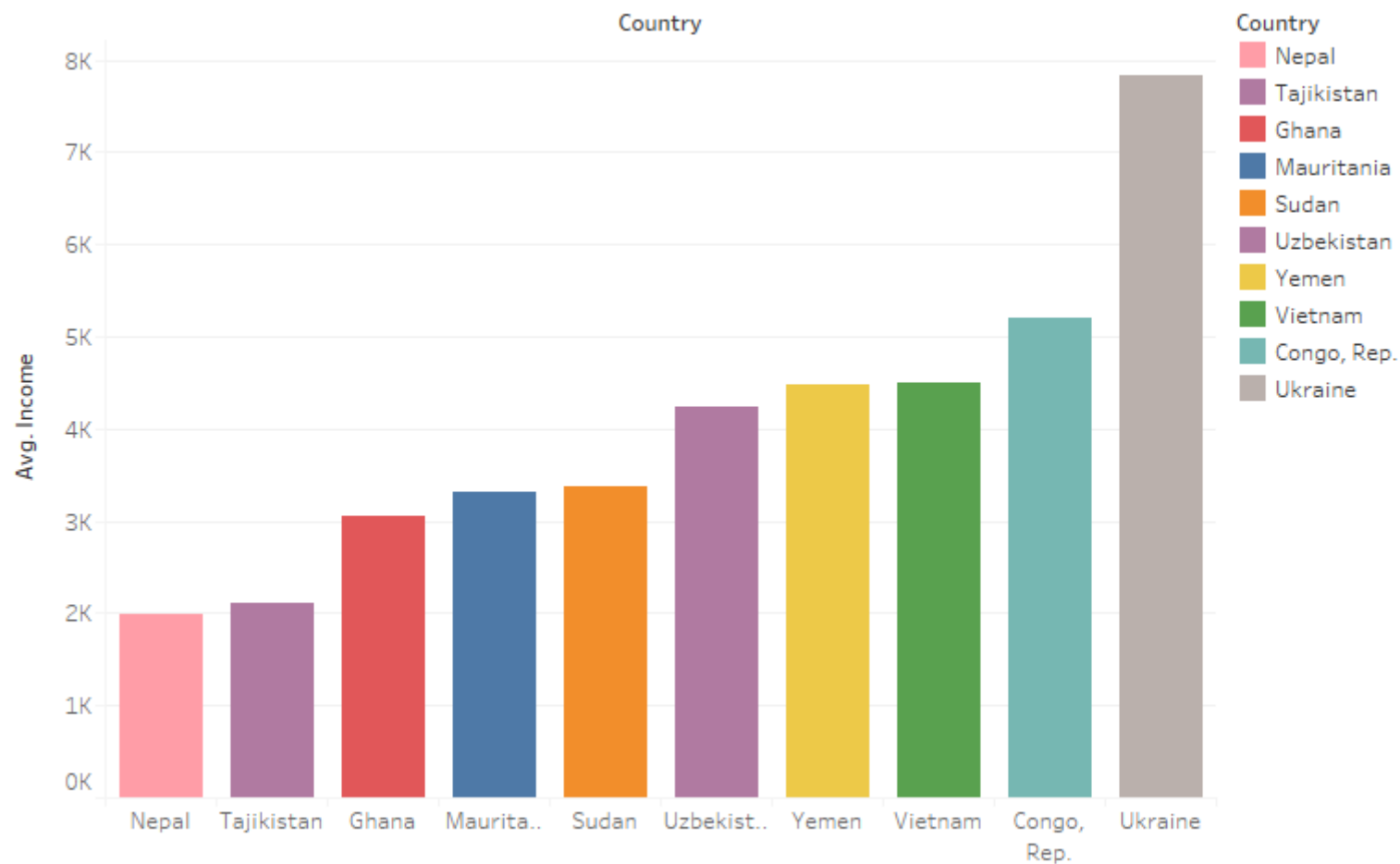


Average of Gdpp for each Country. Color shows details about Country. The view is filtered on Country, which keeps 10 of 25 members.

# <Results>

Plot 3

Sheet 3

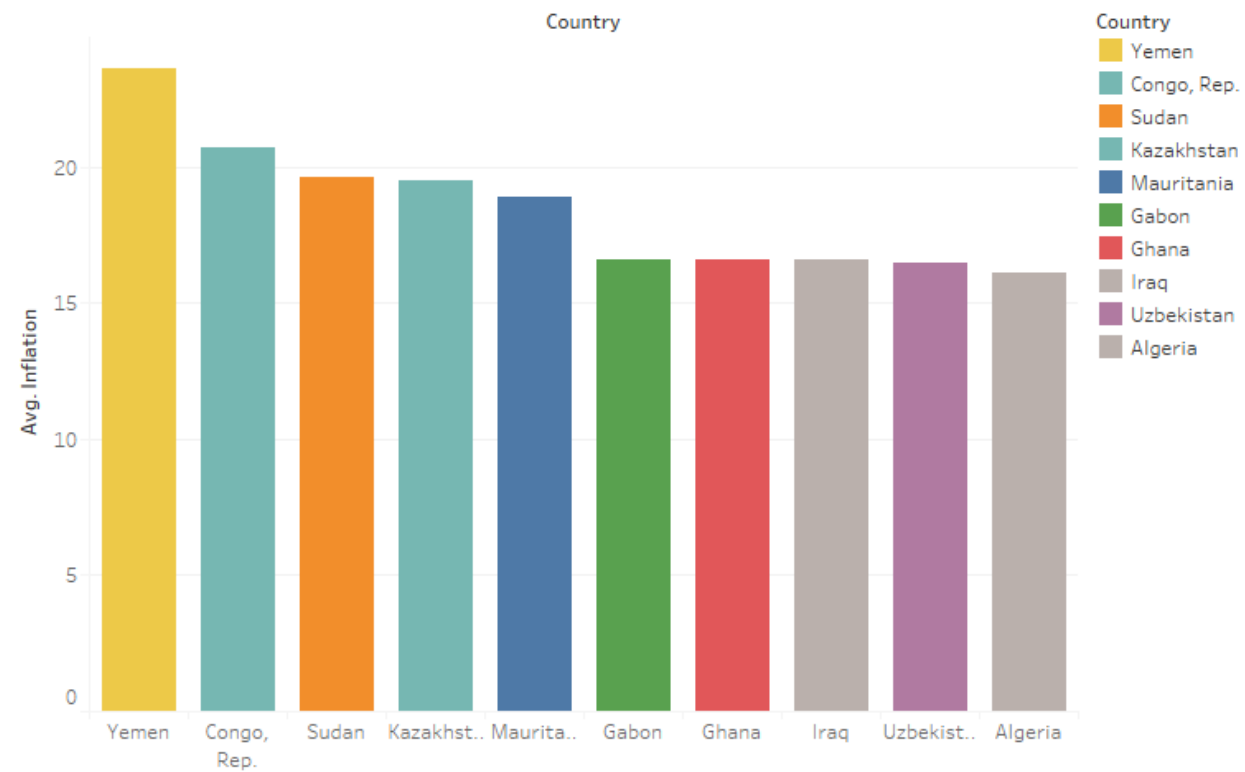


Average of Income for each Country. Color shows details about Country. The view is filtered on Country, which keeps 10 of 25 members.



# <Conclusions>

Sheet 4



Average of Inflation for each Country. Color shows details about Country. The view is filtered on Country, which keeps 10 of 25 members.

<final list of countries>

Top 5 countries for dire need of aid

1.SUDAN

2.YEMEN

3.CONGO REPUBLIC

4.GHANA

5.MAURITANIA