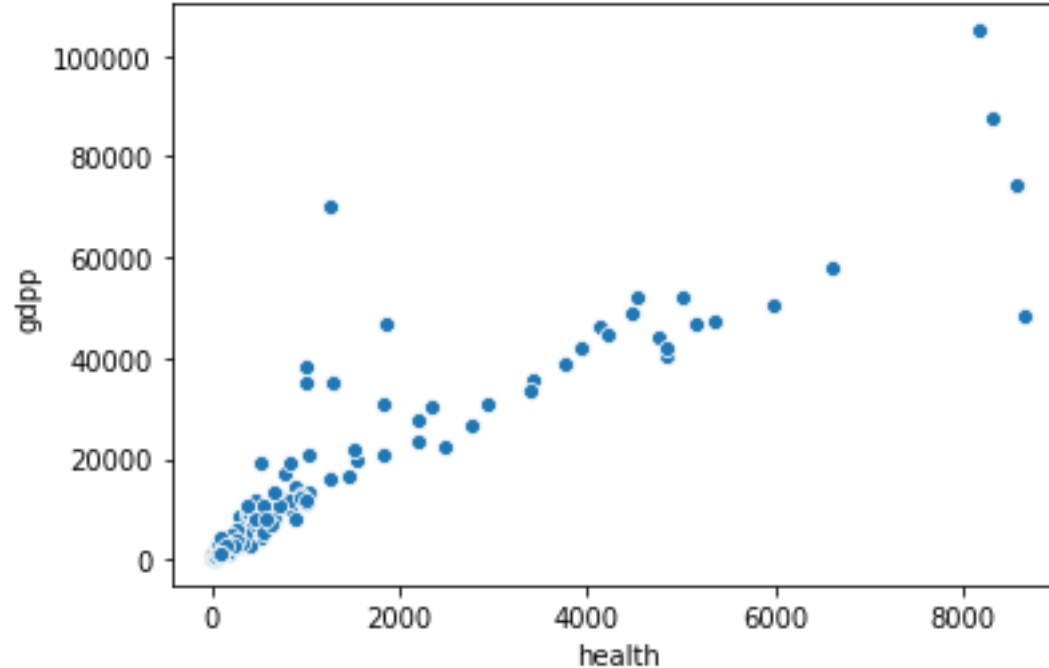


# Clustering Assignment

# Problem Statement:

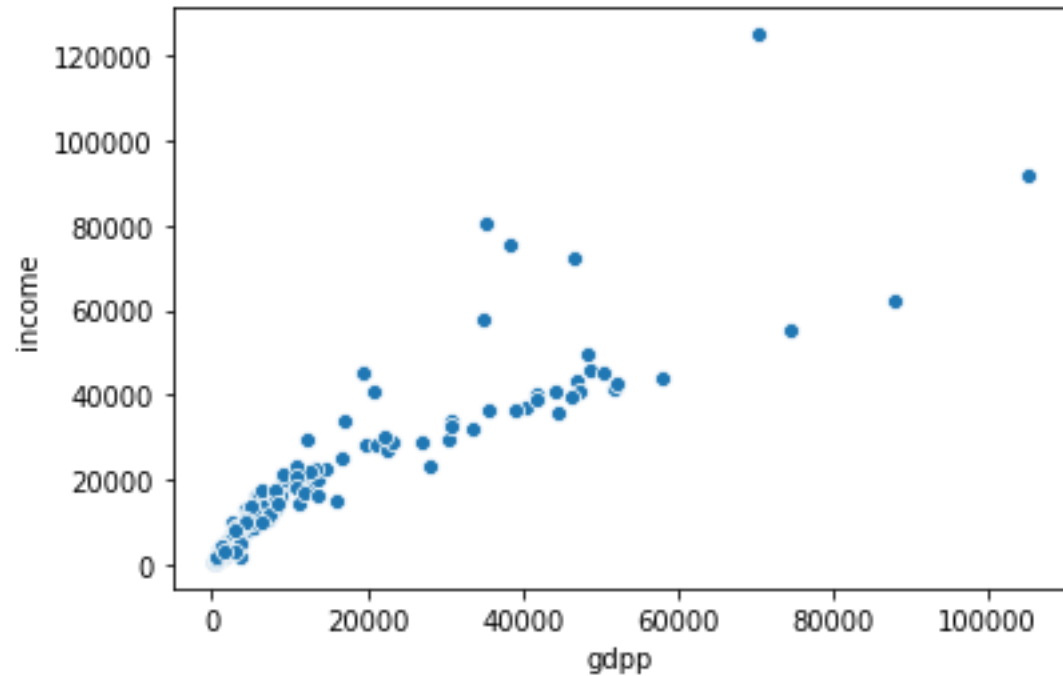
- HELP International is an international humanitarian NGO which has raised around \$ 10 million. The NGO is committed to spend this amount in fighting poverty and providing the people of backward countries with basic amenities and relief during the time of disasters and natural calamities.
- We should categorize the countries using some socio-economic and health factors that determine the overall development of the country. Then you need to suggest the countries which the CEO needs to focus on the most.

# Plot: health vs gdpp:



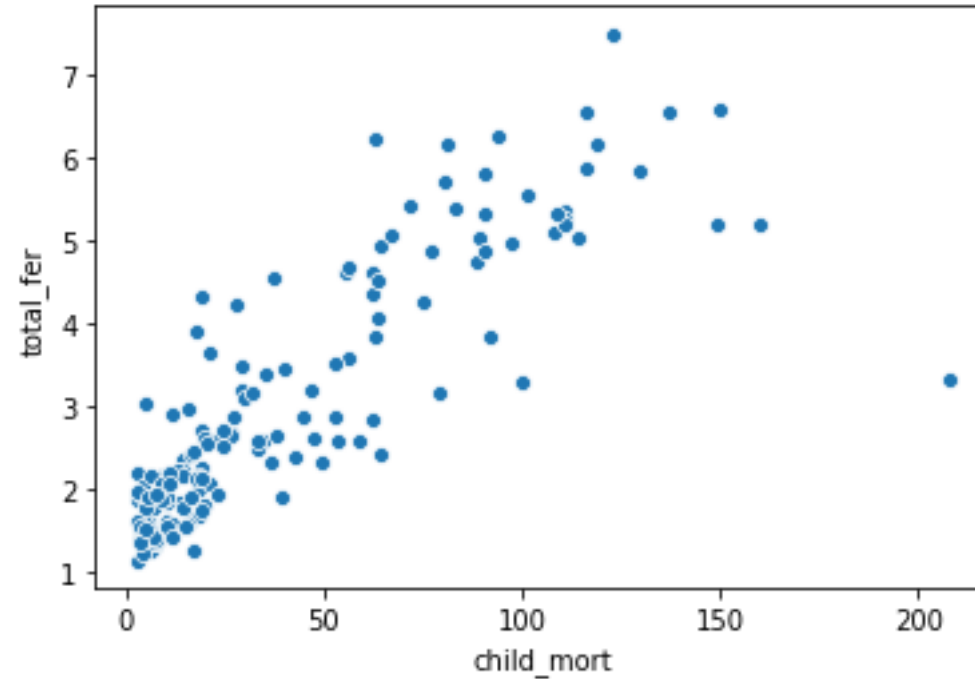
As we can see from the scatter plot, amount spent on ``health`` is relatively high for countries with high ``gdpp`` than the countries with low gdpp.

# Plot: gdpp vs income



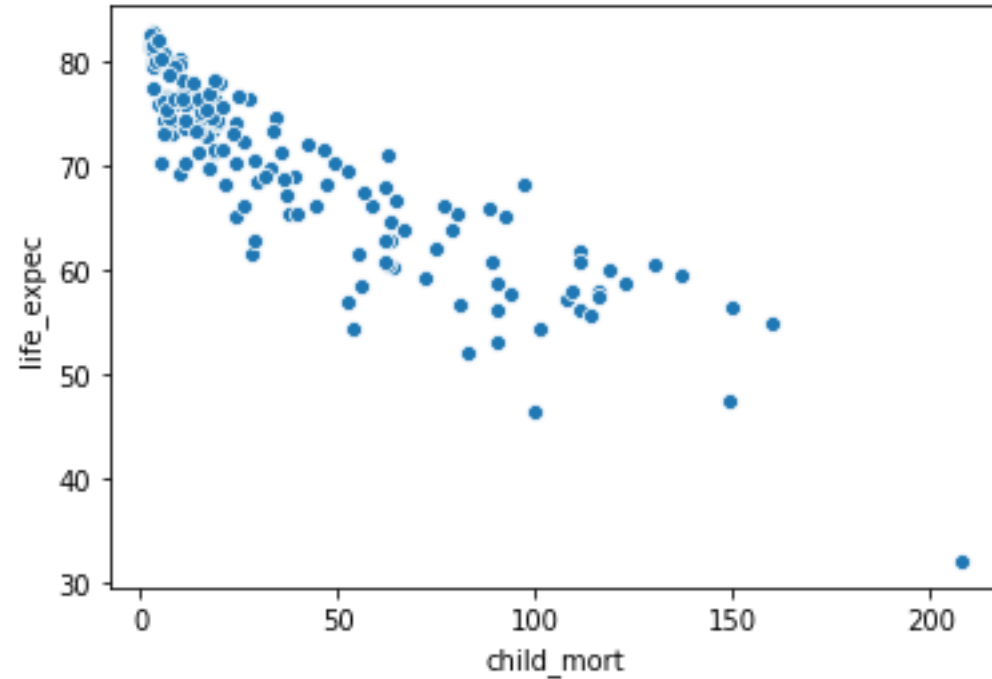
From above scatter plot, net income per person is high for countries with high ``gdpp``.

# Plot: child\_mort vs total\_fer



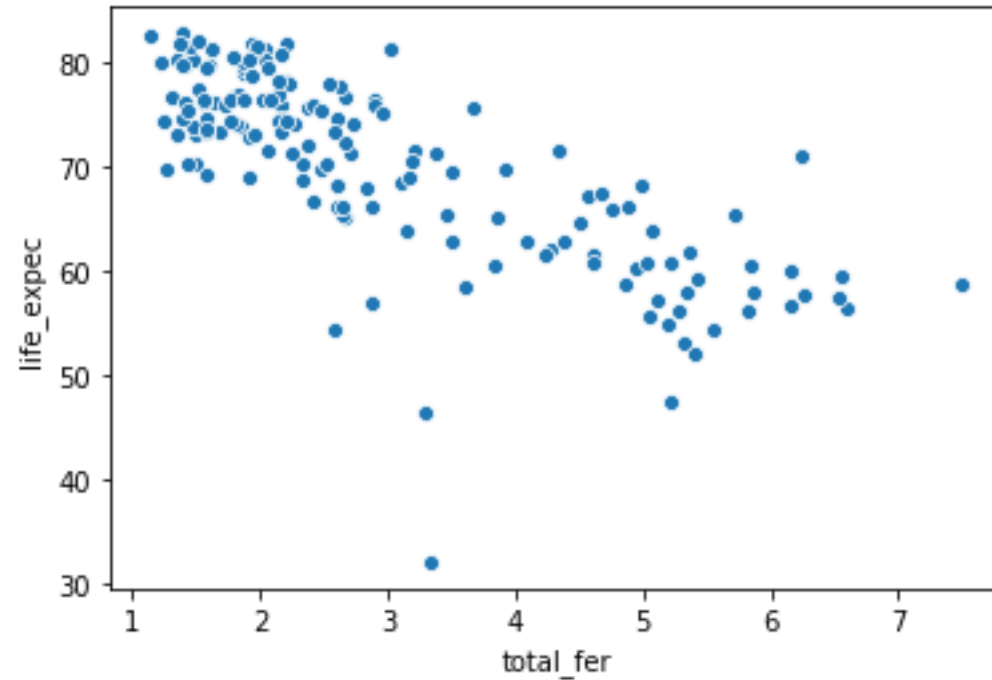
From above plot, ``child\_mort`` and ``total\_fer`` are highly correlated. Deaths of children are increasing with increase in fertility rate.

# Plot: child\_mort vs life\_expec



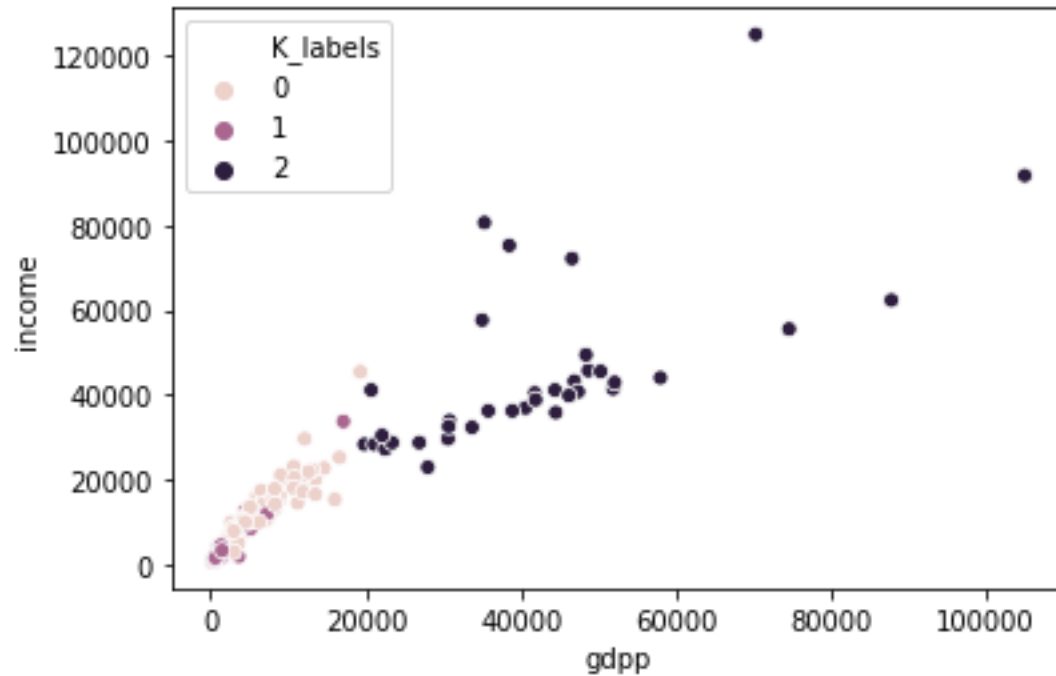
From above plot, ``life\_expec`` for a new born is low for the countries with high ``child\_mort``.

# Plot: total\_fer and life\_expec



From above plot, increase in ``total\_fer`` leads to lower ``life\_expec`` because higher ``total\_fer`` leads to higher ``child\_mort`` which in turn is causing lower ``life\_expec``.

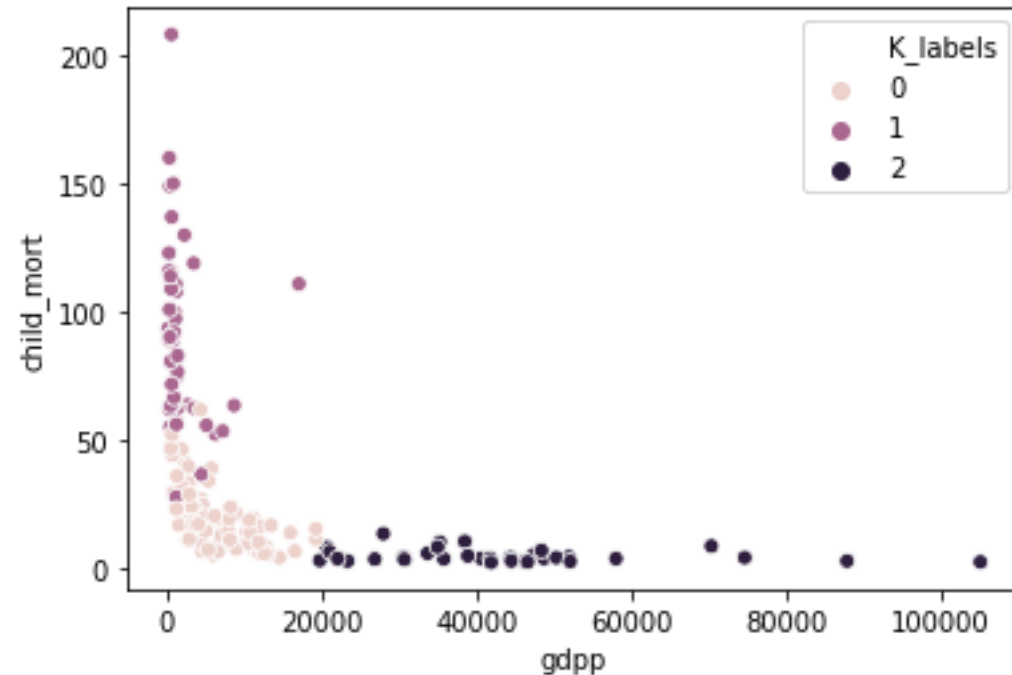
# Cluster Results: gdpp vs income



From above plot, we can say that ``income`` is high for countries with high ``gdpp``.

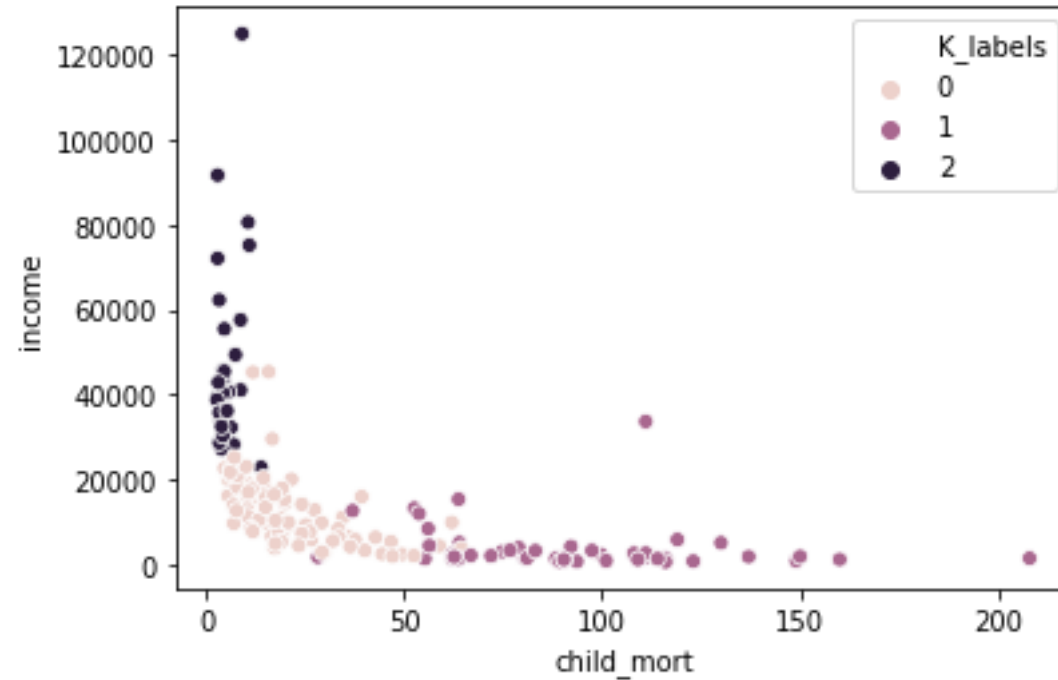


# Cluster Results: gdpp vs child\_mort



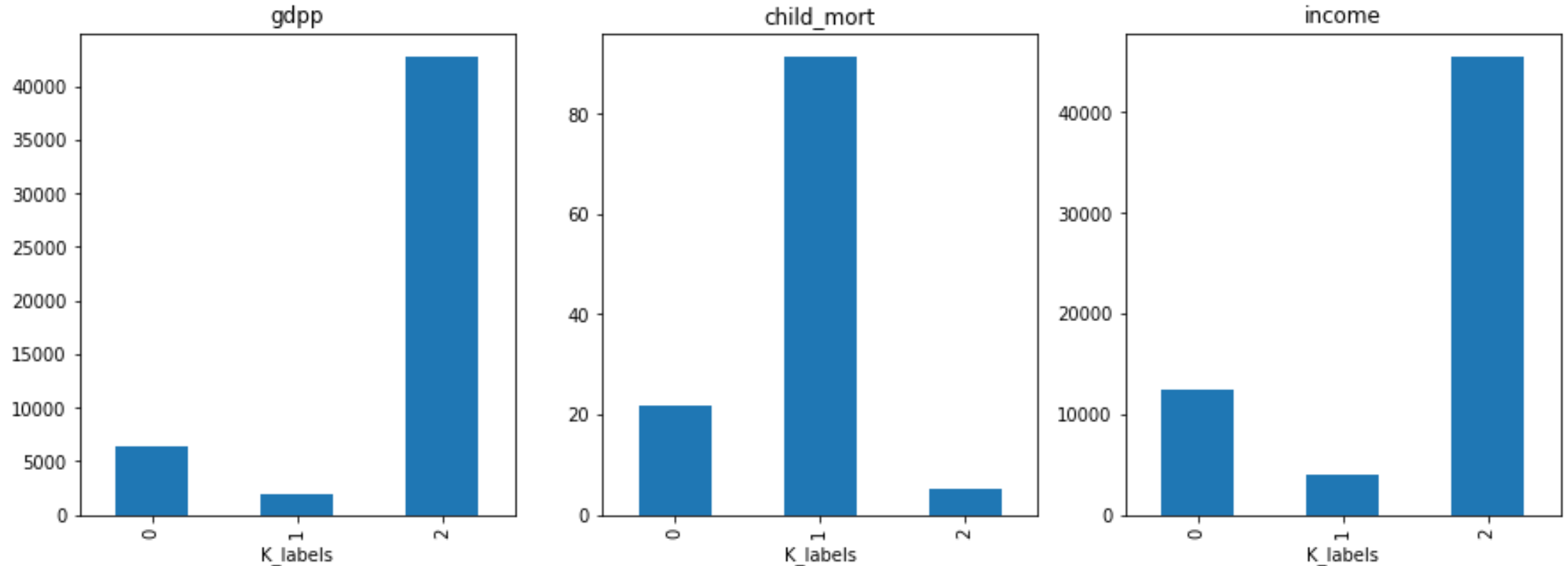
From above plot, ``child\_mort`` is high for countries with low ``gdpp`` because they are not spending enough money on Health due to their economic situation.

# Cluster Results: child\_mort vs income



From above plot, ``child\_mort`` is high for countries with low ``income`` as ``income`` and ``gdpp`` are strongly correlated. Low income means low gdpp which means low spendings on health.

# Average value of gdpp, child\_mort and income



From above plots, ``gdpp`` and ``income`` for cluster 1 are very low and ``child\_mort`` for cluster 1 is almost triple than the second best. So we should target the countries from cluster 1.

# Under-developed countries:

- Below are the countries which fall in cluster 1 i.e. countries that have **low gdpp** and **income** and have **high child\_mort** rate.

Burundi, Liberia, Congo, Dem. Rep., Niger, Sierra Leone, Madagascar, Mozambique, Central African Republic, Malawi, Eritrea, Togo, Guinea-Bissau, Afghanistan, Gambia, Rwanda, Burkina Faso, Uganda, Guinea, Haiti, Tanzania, Mali, Benin, Comoros, Chad, Kenya, Senegal, Pakistan, Lao, Lesotho, Mauritania, "Cote d'Ivoire", Solomon Islands, Cameroon, Ghana, Yemen, Zambia, Sudan, Kiribati, Nigeria, Congo, Rep., Angola, Timor-Leste, Iraq, Namibia, Botswana, South Africa, Gabon, Equatorial Guinea