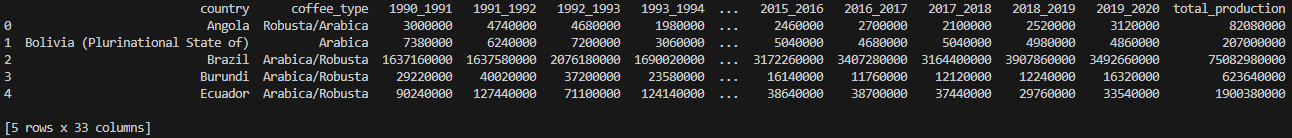
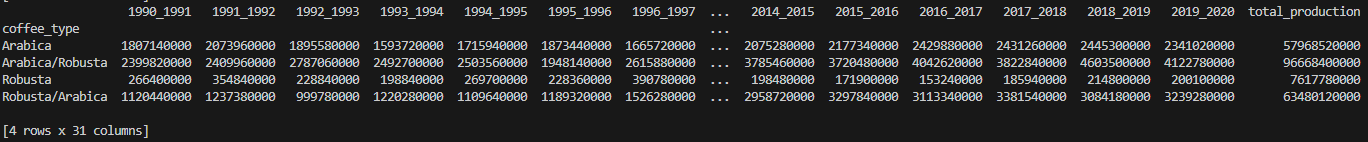
**Coffee Production Analysis**

**Data Aggregation**





* By using groupby[(‘coffee\_type’]) combined with .sum() I am able to aggregate the amount of coffee produced for each coffee type by year.
* I had also dropped the Country column since it is not relevant currently

**Data Transformation**

A black background with white text

AI-generated content may be incorrect.

* To transpose the data, I just used the .transpose() function which easily moved year as the index and type of coffee to columns

**Correlation Analysis**

A screen shot of a computer

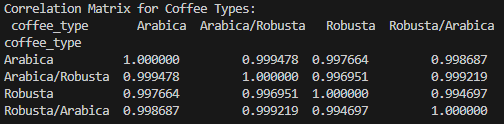
AI-generated content may be incorrect. A screen shot of a computer

AI-generated content may be incorrect.

* For correlation analysis, I used the .corr() function on the cleaned and transposed dataframe which returned the first image above.
* By using .abs().unstack() I am able to retrieve the dataframe as a series with multiple indexes.
* I further cleaned the correlation analysis by removing the self-correlation values
  + 

**Questions**

1. Examine the correlation matrix. Which two coffee types have the **strongest** correlation in production volumes over the years? What might this imply about their production dynamics?
2. Identify the two coffee types with the **weakest** correlation. Discuss possible reasons for this weak relationship and any external factors that might influence these production types differently.





* This is interesting because even though there is a strongest and weakest pair of correlations, all of them are very high at 0.99 on a scale from 0 – 1. The correlations in production years are possibly from the fact that Arabica and Robusta coffee beans are produced in the same regions which could closely align their demands.
* Also, that Arabica and Arabica/Robusta blends can have similar demands due to Arabica still being in the blend. This can also be true for Robusta and Robusta/Arabica blends as well
* Going back to the initial point, there is the “weakest pair” in relation to the strongest pair, but in general it still holds a very strong correlation.