Memo

To: Dr. Landowski

From: Seshu Miriyala

Date: 02/18/2021

Re: Project Proposal

Topic:

Analyze of coffee consumption and production

Data Description

The data set contains 12 csv files which have a broad areas of life cycle of coffee including production, consumption, import, export from 1990 till 1998.

Each file has the following columns.

Field	Description	Sample data
Country	Country of origin of the coffee beans	Austria, Denmark, etc.
1990 – 1998	Production, Consumption of Coffee in thousand bags of 60Kgs each.	1302.0859699, 908.5420009
Price	Price of the coffee in USD/Kg	1.09

Research Questions

1. Which country consumes most of the coffee it produced?
2. Which country has best coffee production?
3. How are farmers paid for the coffee they grow in each country?
4. Which country exports most of the coffee?

Data Preparation

- Load the data into Python
- Concatenate the data from each file into one data set
- Filter the data set to only records in Groups we are interested in analyzing
- Clean the data from missing values and outliers, etc.
- Convert the columns to appropriate dTypes like prices should be in float, countries should be categories, etc.
- Generate the appropriate graphs using matplotlib
- If possible, create a prediction model to predict the coffee consumption or production in 1999.