

Analyzing fertilizers consumption over the in Pakistan

Agriculture is an important sector of Pakistan's economy, accounting for approximately 26 percent of the country gross domestic product (GDP). The aim of this notebook is to analysis the fertilizer uses in Pakistan over the years from 1961 to 2019.

The dataset is taken from **Food and Agriculture Organization FAO**. The organization is a specialized agency of the United Nations leading international efforts to defeat hunger.

About Dataset:

- Domain Code - Regenerative Food Network (RFN)
- Domain - Common nutrients in fertilizer
- Area Code - Country area code in FAO's member list (165)
- Area - Country as a FAO's member (Pakistan)
- Element Code - Data collection type associated with the commodity
- Element - Agricultural data
- Item Code - Item code in FAO list (3102, 3103, 3104)
- Item - Item name (Nitrogen, Phosphate, Potash)
- Year Code - Year code
- Year - Year
- Unit - Unit used to calculate the total amount of a nutrient per year (ton)
- Value - Total amount of a nutrient used per year
- Flag - Qm, X, Fb, W, Fm
- Flag Description - Authorities who collected data

Preparing Tools

We will need to use some libraries to work with the data. Let's import libraries we are going to use.

```
In [ ] :  
# Import libraries  
import numpy as np  
import pandas as pd  
import seaborn as sns  
import plotly.graph_objects as go  
import plotly.express as px  
from plotly.subplots import make_subplots  
import matplotlib.pyplot as plt  
%matplotlib inline
```

Load data

We have data in csv file `fertilizers_pk.csv`, let's import data in dataframe.

```
In [ ] :  
# Load dataset  
data_pak = pd.read_csv("../datasets/fertilizers_pk.csv")  
data_pak.head()
```

```
Out[ ] :  


|   | Domain Code | Domain                  | Area Code | Area     | Element Code | Element          | Item Code | Item                            | Year Code | Year | Unit   | Value | Flag | Flag Description                                    |
|---|-------------|-------------------------|-----------|----------|--------------|------------------|-----------|---------------------------------|-----------|------|--------|-------|------|-----------------------------------------------------|
| 0 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total)     | 1961      | 1961 | tonnes | 41659 | Qm   | Official data from questionnaires and/or nationa... |
| 1 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1961      | 1961 | tonnes | 500   | Qm   | Official data from questionnaires and/or nationa... |
| 2 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1962      | 1962 | tonnes | 210   | Qm   | Official data from questionnaires and/or nationa... |
| 3 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total)     | 1962      | 1962 | tonnes | 41160 | Qm   | Official data from questionnaires and/or nationa... |
| 4 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total)     | 1963      | 1963 | tonnes | 67620 | Qm   | Official data from questionnaires and/or nationa... |


```

```
In [ ] :  
# Find number of rows and columns in the dataset  
data_pak.shape
```

```
Out[ ] :  
(172, 14)
```

We have 172 observations across 14 features.

Now, let's check if the data has any missing values.

```
In [ ] :  
# Find missing values in each columns  
data_pak.isnull().sum()
```

```
Out[ ] :  
Domain Code      0  
Domain           0  
Area Code        0  
Area             0  
Element Code     0  
Element          0  
Item Code        0  
Item             0  
Year Code        0  
Year             0  
Unit             0  
Value           0  
Flag             0  
Flag Description 0  
dtype: int64
```

There is no missing values in our data which is a good thing.

How about getting some information data?

```
In [ ] :  
# Data information  
data_pak.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 172 entries, 0 to 171  
Data columns (total 14 columns):  
#   Column      Non-Null Count  Dtype  ---  
0   Domain Code  172 non-null    object  
1   Domain      172 non-null    object  
2   Area Code   172 non-null    int64  
3   Area        172 non-null    object  
4   Element Code 172 non-null    int64  
5   Element     172 non-null    object  
6   Item Code   172 non-null    int64  
7   Item        172 non-null    object  
8   Year Code   172 non-null    int64  
9   Year        172 non-null    int64  
10  Unit        172 non-null    object  
11  Value       172 non-null    int64  
12  Flag        172 non-null    object  
13  Flag Description 172 non-null    object  
dtypes: int64(6), object(8)  
memory usage: 18.9+ KB
```

After looking at the data we can say the every column has a right data type.

Let's find out what are the names of nutrient we have in the data so we can analyze them for later use.

```
In [ ] :  
# Get the list of nutrients name  
list(data_pak["Item"].unique())
```

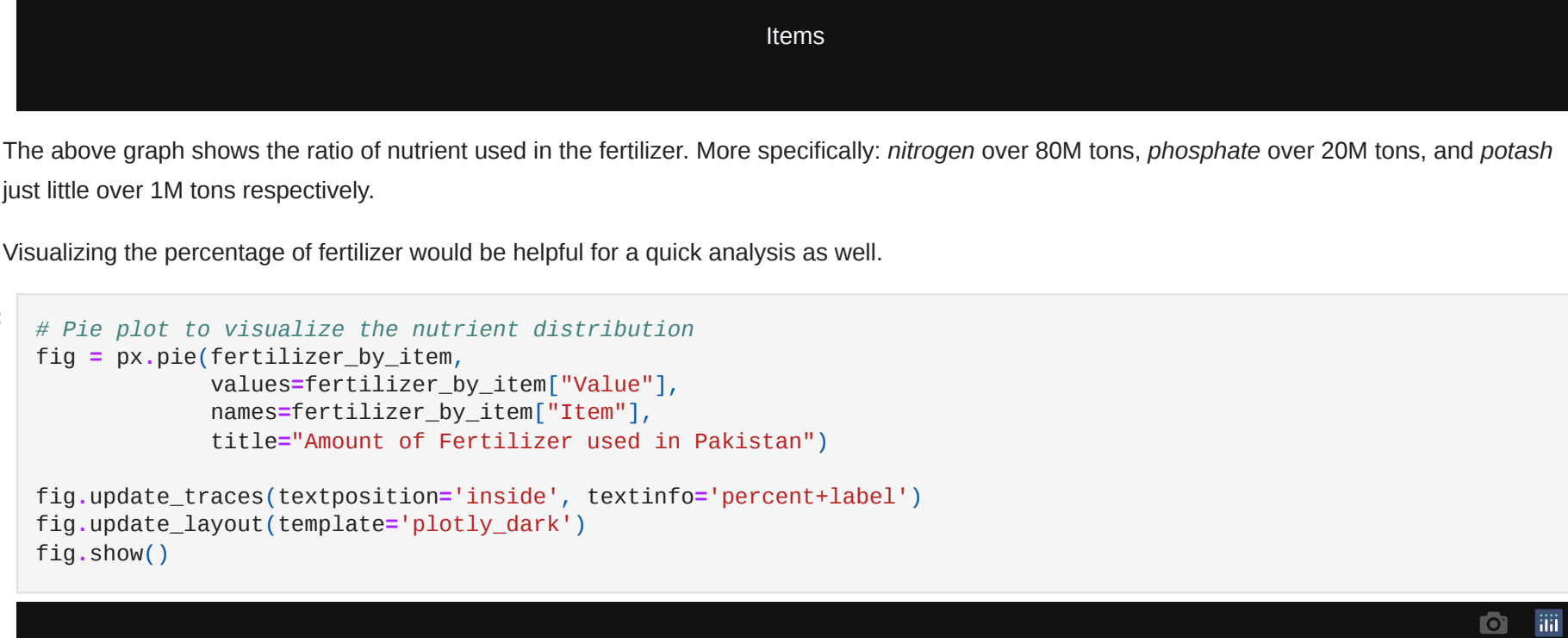
```
Out[ ] :  
['Nutrient nitrogen N (total)',  
 'Nutrient phosphate P2O5 (total)',  
 'Nutrient potash K2O (total)']
```

We can group the value so that we can clearly visualize how much of these three nutrients have been used in Pakistan from 1961 till 2019.

```
In [ ] :  
# Group together "Item" with "Value"  
fertilizer_by_item = data_pak.groupby(["Item"])["Value"].sum().reset_index().sort_values("Value",ascending=False).reset_index()
```

Since, we have the new dataframe `fertilizer_by_item`, we can visualize it.

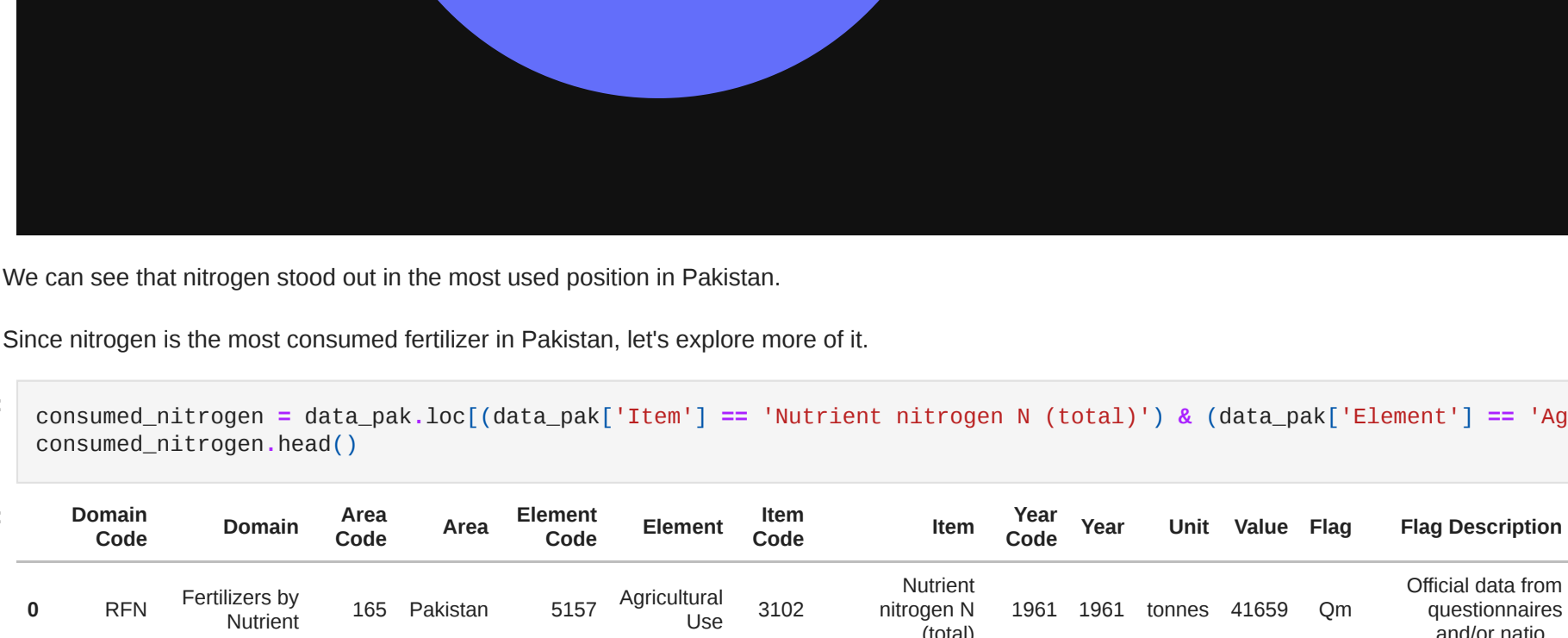
```
In [ ] :  
# Bar plot of the fertilizer  
fig = go.Figure(data=[go.Bar(x=fertilizer_by_item['Item'][:3],  
                             y=fertilizer_by_item['Value'][:3],  
                             text=fertilizer_by_item['Value'][:3],  
                             textposition='auto',  
                             marker_color='red')])  
  
fig.update_layout(title='Amount of fertilizers used in Pakistan since 1961',  
                  xaxis_title='Items',  
                  yaxis_title='Value',  
                  template='plotly_dark')  
  
fig.show()
```



The above graph shows the ratio of nutrient used in the fertilizer. More specifically: *nitrogen* over 80M tons, *phosphate* over 20M tons, and *potash* just little over 1M tons respectively.

Visualizing the percentage of fertilizer would be helpful for a quick analysis as well.

```
In [ ] :  
# Pie plot to visualize the nutrient distribution  
fig = px.pie(fertilizer_by_item,  
            values=fertilizer_by_item["Value"],  
            names=fertilizer_by_item["Item"],  
            title='Amount of Fertilizer used in Pakistan')  
  
fig.update_traces(textposition='inside', textinfo='percent+label')  
fig.update_layout(template='plotly_dark')  
fig.show()
```



We can see that nitrogen stood out in the most used position in Pakistan.

Since nitrogen is the most consumed fertilizer in Pakistan, let's explore more of it.

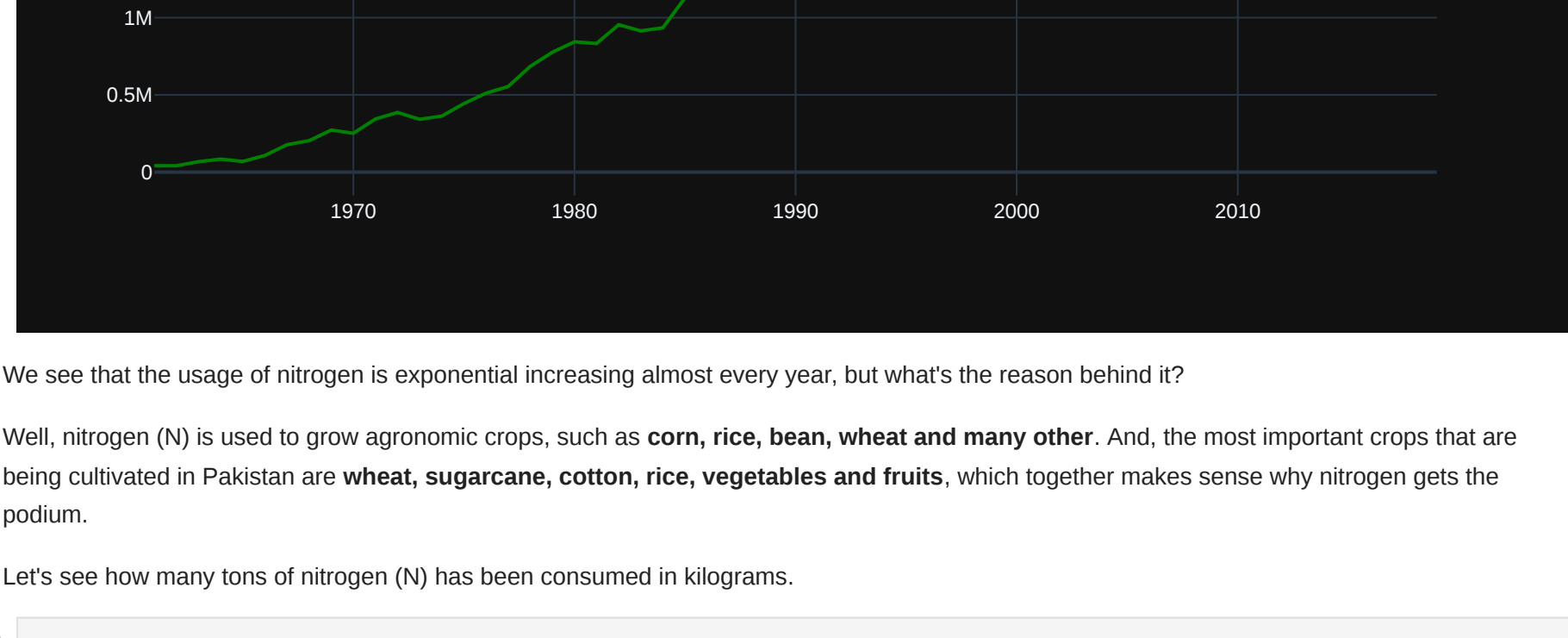
```
In [ ] :  
consumed_nitrogen = data_pak.loc[(data_pak['Item'] == 'Nutrient nitrogen N (total)') & (data_pak['Element'] == 'Agricultural Use')]  
consumed_nitrogen.head()
```

```
Out[ ] :  


|   | Domain Code | Domain                  | Area Code | Area     | Element Code | Element          | Item Code | Item                        | Year Code | Year | Unit   | Value | Flag | Flag Description                                    |
|---|-------------|-------------------------|-----------|----------|--------------|------------------|-----------|-----------------------------|-----------|------|--------|-------|------|-----------------------------------------------------|
| 0 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total) | 1961      | 1961 | tonnes | 41659 | Qm   | Official data from questionnaires and/or nationa... |
| 3 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total) | 1962      | 1962 | tonnes | 41160 | Qm   | Official data from questionnaires and/or nationa... |
| 4 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total) | 1963      | 1963 | tonnes | 67620 | Qm   | Official data from questionnaires and/or nationa... |
| 7 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total) | 1964      | 1964 | tonnes | 84147 | Qm   | Official data from questionnaires and/or nationa... |
| 8 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3102      | Nutrient nitrogen N (total) | 1965      | 1965 | tonnes | 69242 | Qm   | Official data from questionnaires and/or nationa... |


```

```
In [ ] :  
fig = go.Figure()  
  
fig.add_trace(go.Scatter(x=consumed_nitrogen['Year'],  
                        y=consumed_nitrogen['Value'],  
                        mode='lines',  
                        name='Nutrient nitrogen (N)', marker_color='green'))  
  
fig.update_layout(title='Nitrogen consumption over the years in Pakistan',  
                  template='plotly_dark')  
  
fig.show()
```



We see that the usage of nitrogen is exponential increasing almost every year, but what's the reason behind it?

Well, nitrogen (N) is used to grow agronomic crops, such as **corn, rice, bean, wheat and many other**. And, the most important crops that are being cultivated in Pakistan are **wheat, sugarcane, cotton, rice, vegetables and fruits**, which together makes sense why nitrogen gets the podium.

Let's see how many tons of nitrogen (N) has been consumed in kilograms.

```
In [ ] :  
consumed_nitrogen_in_kg = consumed_nitrogen['Value'].sum()  
print(f'Total amount of Nitrogen since 1961: {consumed_nitrogen_in_kg*1000} kilograms')
```

Total amount of Nitrogen since 1961: 94314067000 kilograms

The second most used fertilizer in Pakistan is phosphate (P2O5). Let's compare them.

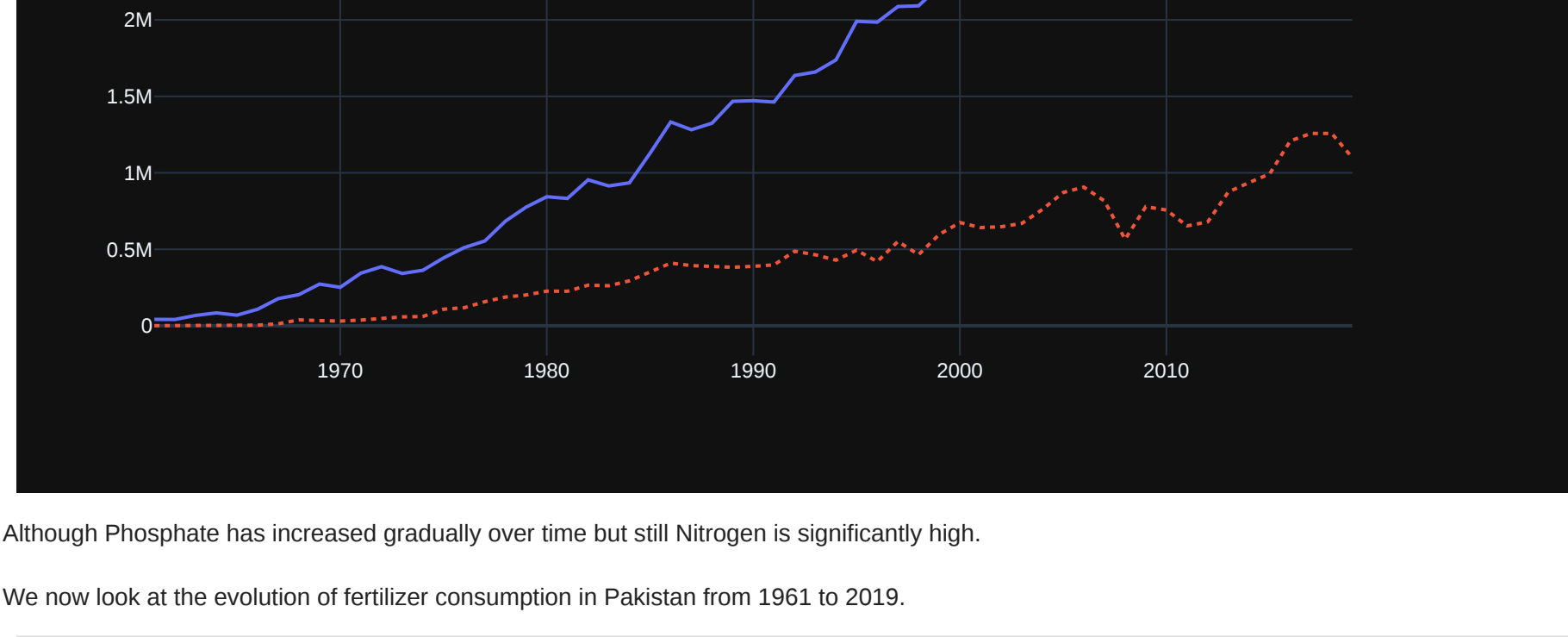
```
In [ ] :  
consumed_phosphate = data_pak.loc[(data_pak['Item'] == 'Nutrient phosphate P2O5 (total)') & (data_pak['Element'] == 'Agricultural Use')]  
consumed_phosphate.head()
```

```
Out[ ] :  


|   | Domain Code | Domain                  | Area Code | Area     | Element Code | Element          | Item Code | Item                            | Year Code | Year | Unit   | Value | Flag | Flag Description                                    |
|---|-------------|-------------------------|-----------|----------|--------------|------------------|-----------|---------------------------------|-----------|------|--------|-------|------|-----------------------------------------------------|
| 1 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1961      | 1961 | tonnes | 500   | Qm   | Official data from questionnaires and/or nationa... |
| 2 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1962      | 1962 | tonnes | 210   | Qm   | Official data from questionnaires and/or nationa... |
| 5 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1963      | 1963 | tonnes | 630   | Qm   | Official data from questionnaires and/or nationa... |
| 6 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1964      | 1964 | tonnes | 1029  | Qm   | Official data from questionnaires and/or nationa... |
| 9 | RFN         | Fertilizers by Nutrient | 165       | Pakistan | 5157         | Agricultural Use | 3103      | Nutrient phosphate P2O5 (total) | 1965      | 1965 | tonnes | 1245  | Qm   | Official data from questionnaires and/or nationa... |


```

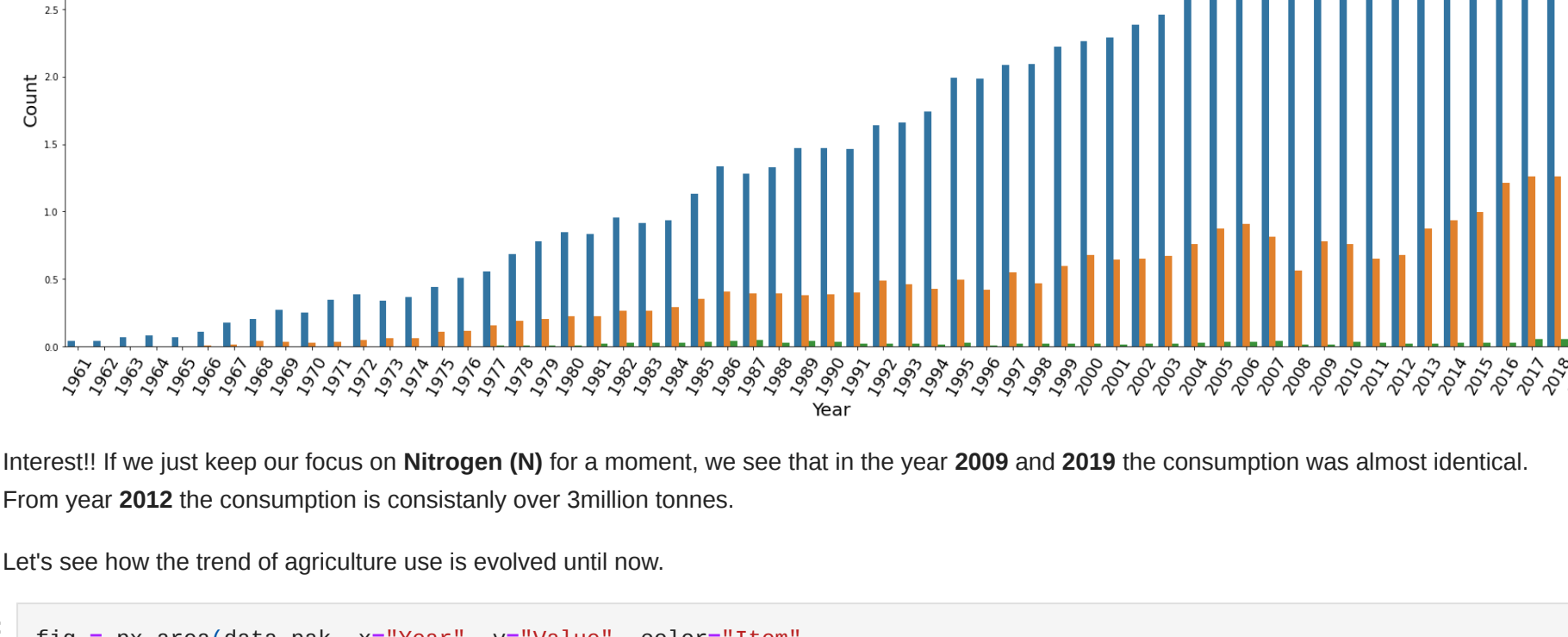
```
In [ ] :  
fig = go.Figure()  
fig.add_trace(go.Scatter(x=consumed_nitrogen['Year'],  
                        y=consumed_nitrogen['Value'],  
                        mode='lines',  
                        name='Nitrogen'))  
fig.add_trace(go.Scatter(x=consumed_phosphate['Year'],  
                        y=consumed_phosphate['Value'],  
                        mode='lines',  
                        name='Phosphate',  
                        line=dict(dash='dot')))  
  
fig.update_layout(title='Comparison between Nitrogen (N) and Phosphate (P2O5)',  
                  template='plotly_dark')  
  
fig.show()
```



Although Phosphate has increased gradually over time but still Nitrogen is significantly high.

We now look at the evolution of fertilizer consumption in Pakistan from 1961 to 2019.

```
In [ ] :  
# Set figure  
plt.subplots(figsize=(30, 10))  
sns.set_context("notebook", font_scale=1.5, rc={"lines.linewidth": 2})  
  
# Draw bar plot  
sns.barplot(x="Year", y="Value", data=data_pak, hue="Item")  
plt.xlabel(xlabel='Year', size=20)  
plt.xticks(rotation=60, size=8)  
plt.ylabel(ylabel='Count', size=20)  
plt.title('Fertilizer consumption trend in Pakistan since 1961', size=30)  
plt.legend()  
plt.show()
```



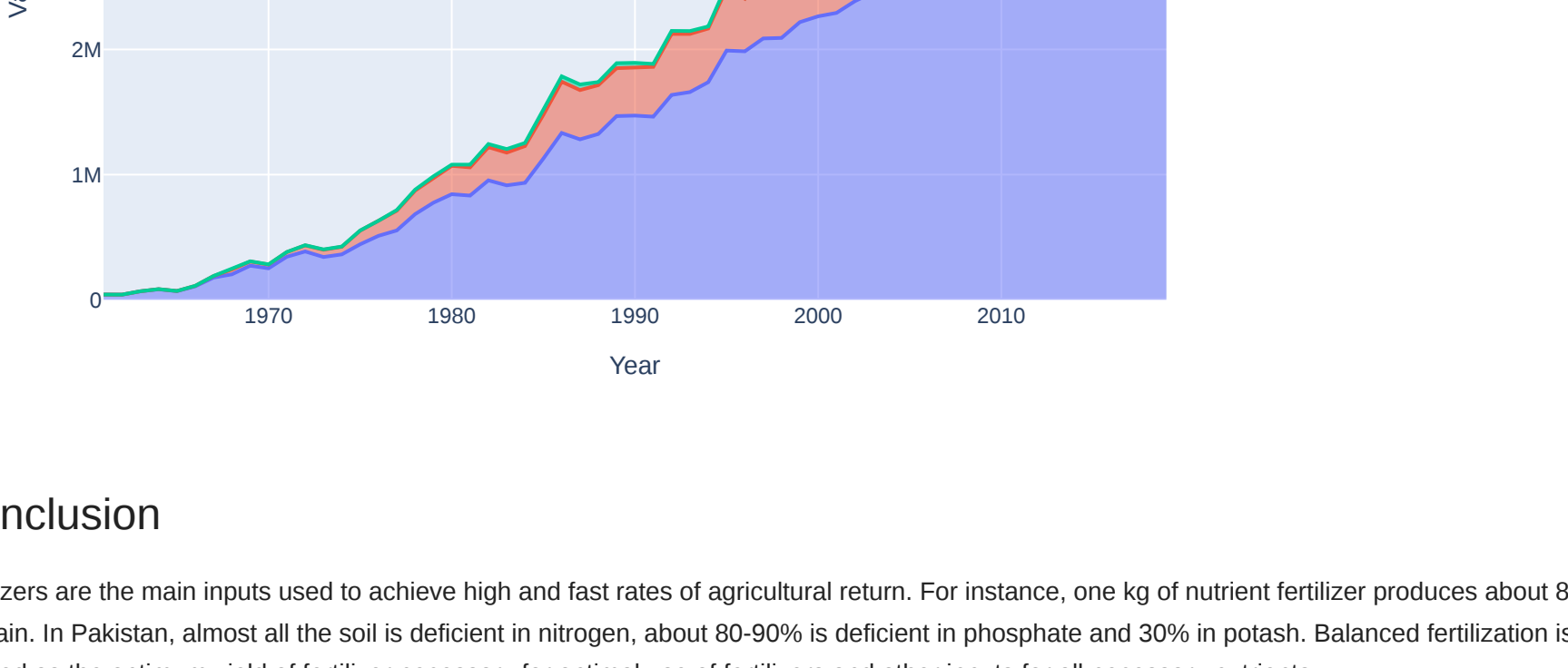
Interest!! If we just keep our focus on **Nitrogen (N)** for a moment, we see that in the year **2009** and **2019** the consumption was almost identical.

From year **2012** the consumption is consistently over 3million tonnes.

Let's see how the trend of agriculture use is evolved until now.

```
In [ ] :  
fig = px.area(data_pak, x="Year", y="Value", color="Item",  
             line_group="Item", title="Consumption of fertilizers in Pakistan")  
fig.show()
```

Consumption of fertilizers in Pakistan



Conclusion

Fertilizers are the main inputs used to achieve high and fast rates of agricultural return. For instance, one kg of nutrient fertilizer produces about 8 kg of grain. In Pakistan, almost all the soil is deficient in nitrogen, about 80-90% is deficient in phosphate and 30% in potash. Balanced fertilization is defined as the optimum yield of fertilizer necessary for optimal use of fertilizers and other inputs for all necessary nutrients.

Pakistan has established the importance of fertilizers in increasing agricultural production as a developing country. Fertilizers are commonly believed to be as important, and contribute up to 50% of the growth in output.

The agricultural sector of Pakistan remains the main economic sector, and contributes to about 60% of the country's GDP. Production in the agricultural sector is the main revenue source for the overall economy in Pakistan. With time, the share of the agricultural sector declined significantly due to technical development.