iran_agriculture_data_analysis

May 28, 2020

```python
[1]: import pandas as pd
     import seaborn as sns
     import numpy as np
     import matplotlib.pyplot as plt
```

```python
[2]: iran_price = pd.read_csv('dataset/FAOSTAT_IRAN_PRODUCER_PRICES.csv')
     iran_exchange = pd.read_csv('dataset/FAOSTAT_IRAN_EXCHANGE_RATES_ANNUAL.csv')
     iran_population  = pd.read_csv('dataset/FAOSTAT_IRAN_POPULATION.csv')
```

```python
[3]: iran_price['Year'].min()
```

```
[3]: 1991
```

```python
[4]: iran_price = iran_price.drop(['Domain Code' , 'Domain' , 'Area Code' , 'Area' ,␣
     ↪'Element Code' ,
                                     'Item Code' , 'Year Code','Flag' , 'Flag␣
     ↪Description'] , axis = 1)
```

```python
[5]: # product that every family needs
     home_basket = ['Wheat', "Meat live weight, cattle", "Meat live weight, sheep",␣
     ↪"Milk, whole fresh cow", "Potatoes", "Tomatoes", "Onions, dry",
     "Honey, natural", "Tea", "Watermelons", "Sunflower seed", "Soybeans", "Apples",␣
     ↪"Rice, paddy", "Eggplants (aubergines)", "Garlic",
     "Fruit, citrus nes", "Fruit, fresh nes", "Vegetables, leguminous nes", "Beans,␣
     ↪dry", "Dates", "Lentils", "Oranges", "Tobacco, unmanufactured" , 'Lemons and␣
     ↪limes']
```

```python
[6]: # mask home basket for data
     for row, col in iran_price.iterrows():
         if col['Item'] not in home_basket:
             iran_price.drop(row, axis= 0, inplace= True)
     iran_price.reset_index(inplace=True)
     del iran_price['index']
```

```python
[7]: iran_exchange = iran_exchange.drop(['Domain Code' , 'Domain' , 'Area Code' ,␣
     ↪'Area' , 'ISO Currency Code' ,
```

```
                                        'Currency' ,'Item' , 'Item Code' , 'Year␣
    ↪Code', 'Note',  'Unit' ,
                                        'Flag' , 'Flag Description'] , axis = 1)
```

```
[8]: iran_exchange['Value'] = iran_exchange['Value'].astype('int64')
     new_exchange = {'Year': [2019,2020] , 'Value': [140000 , 180000] }
     new_exchange = pd.DataFrame(new_exchange)
     iran_exchange = iran_exchange.append(new_exchange)
     iran_exchange.reset_index(inplace=True)
     del iran_exchange['index']
```

```
[9]: iran_price = iran_price.merge(iran_exchange, left_on= "Year", right_on="Year",␣
     ↪how= "left", suffixes=(" in Dollars", " Dollar per Rials"))
     iran_price['Value in Rials'] = iran_price["Value in Dollars"] *␣
     ↪iran_price["Value Dollar per Rials"]
     del iran_price['Value Dollar per Rials']
```

```
[10]: f, ax = plt.subplots(figsize=(16, 7))
      mask = np.zeros_like(iran_price.corr(), dtype=np.bool)
      mask[np.triu_indices_from(mask)] = True

      plt.title('Pearson Correlation Matrix',fontsize=12)

      sns.heatmap(iran_price.corr(),linewidths=0.25,vmax=0.7,square=True,cmap="BuGn",
                  linecolor='w',annot=True,annot_kws={"size":
      ↪15},mask=mask,cbar_kws={"shrink": .9});
```

Pearson Correlation Matrix
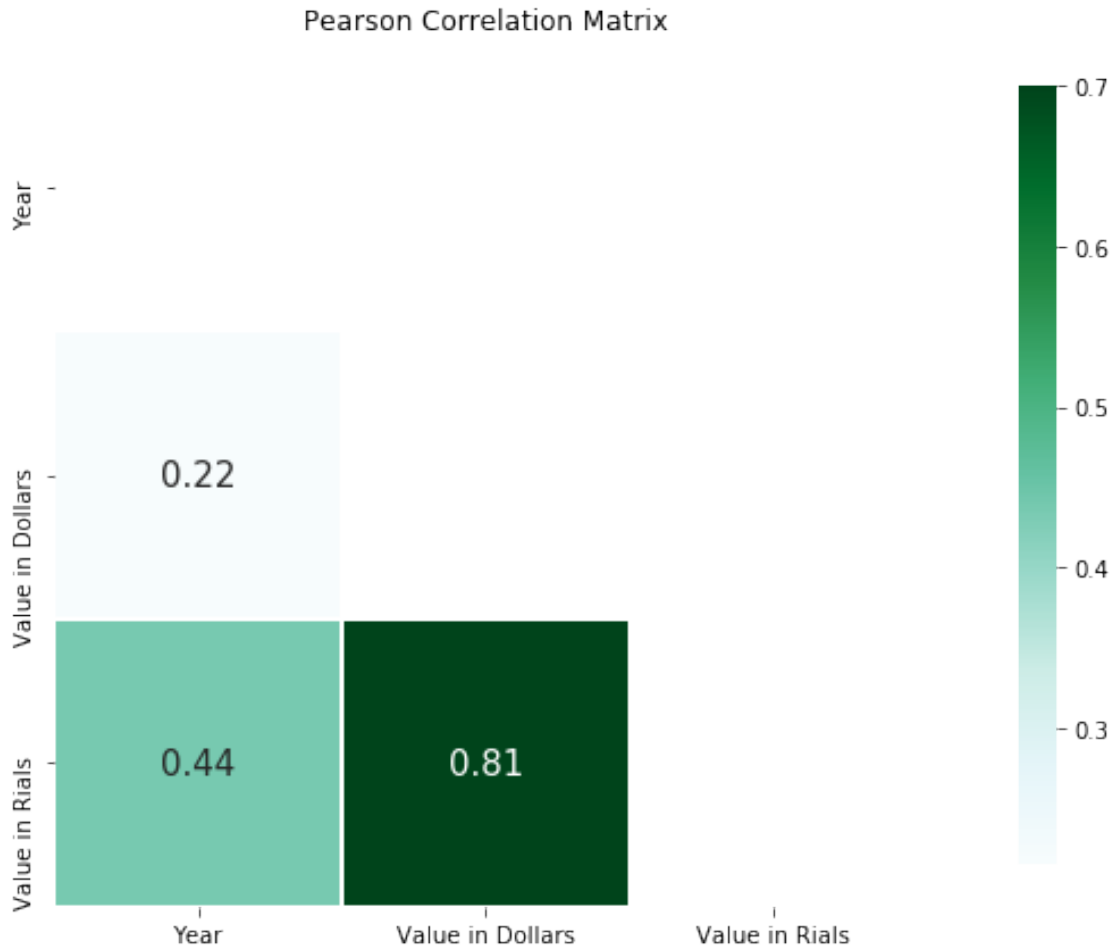

```
[11]: iran_price['Price per Kg in Rial'] = iran_price['Value in Rials'] / 1000
      del iran_price['Value in Dollars']
      del iran_price['Value in Rials']
      del iran_price['Element']
```

```
[12]: iran_price['President'] = 0
      for i in range(len(iran_price)):
          if iran_price.loc[i, 'Year'] >= 1991 and iran_price['Year'].iloc[i] < 1997:
              iran_price.loc[i, 'President'] = 'Akbar Hashemi Rafsanjani'
          elif iran_price.loc[i, 'Year'] >= 1997 and iran_price['Year'].iloc[i] <↵
      ↪2005:
              iran_price.loc[i, 'President'] = 'Mohammad Khatami'
          elif iran_price.loc[i, 'Year'] >= 2005 and iran_price['Year'].iloc[i] <↵
      ↪2013:
              iran_price.loc[i,'President'] = 'Mahmoud Ahmadinejad'
          elif iran_price.loc[i, 'Year'] >= 2013 and iran_price['Year'].iloc[i] <=↵
      ↪2018:
```
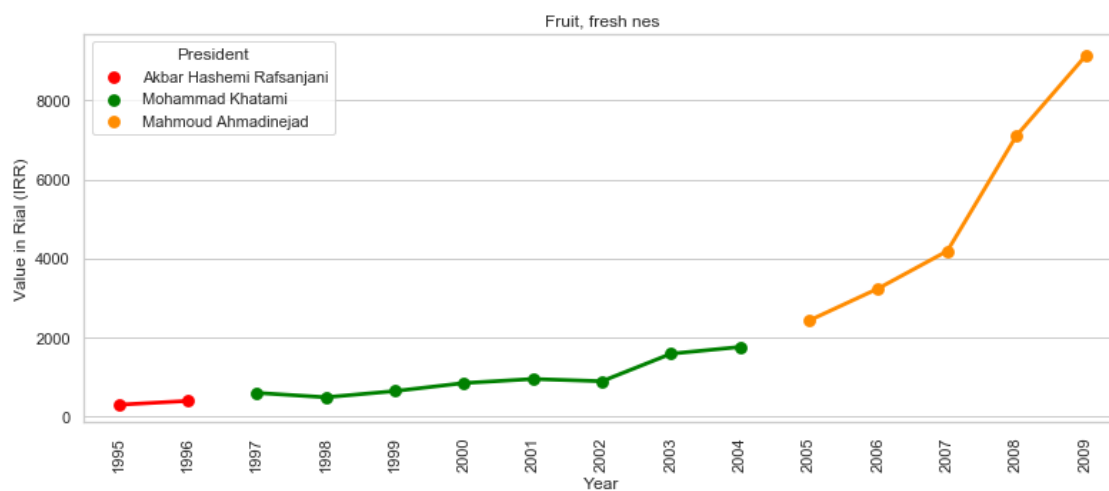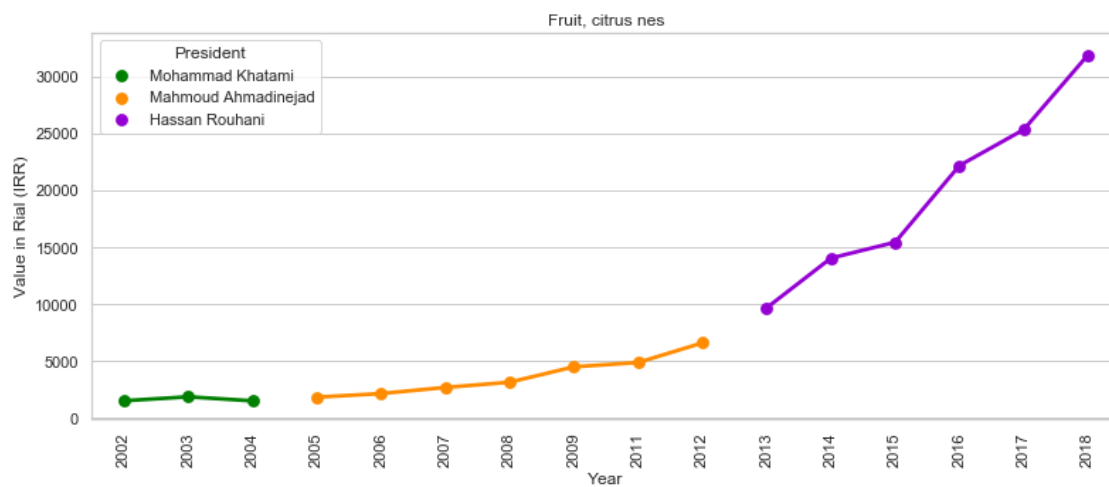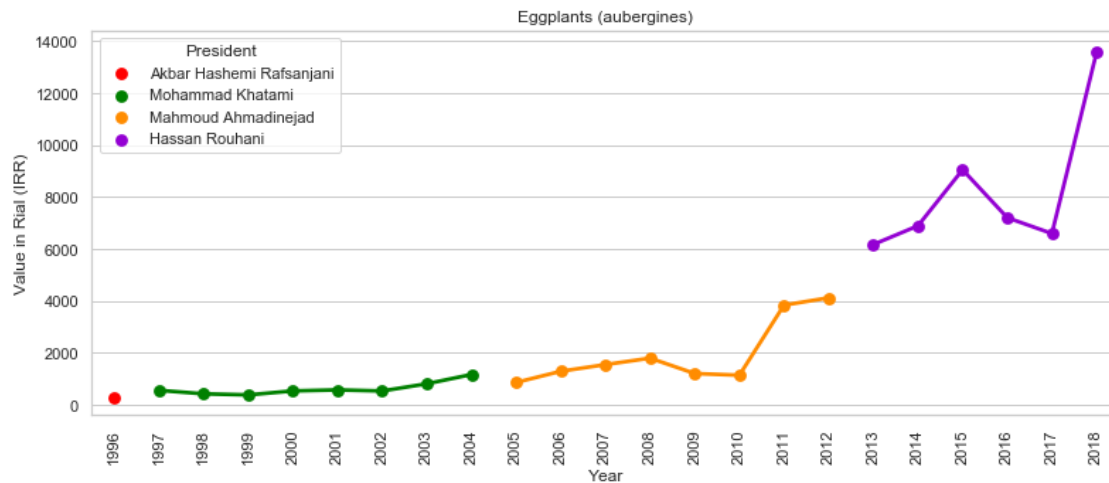
```
              iran_price.loc[i, 'President'] = 'Hassan Rouhani'
```

```
[13]: iran_population = iran_population[iran_population['Element'] == 'Total␣
      ↪Population - Both sexes']
      iran_population = iran_population.drop(['Domain Code' , 'Domain' , 'Area Code'␣
      ↪, 'Area' , 'Element Code' ,'Element' ,'Item' , 'Item Code' , 'Year Code'
                ,'Note',  'Unit'  , 'Flag' , 'Flag Description'] , axis = 1)
      iran_population = iran_population[iran_population['Year'] >= 1991]
```

```
[14]: iran_price = iran_price.merge(iran_population, left_on= "Year",␣
      ↪right_on="Year", how= "left", suffixes=(" in Dollars", " Dollar per Rials"))
      iran_price.rename(columns={'Value': 'Population'}, inplace=True)
      iran_price['Population'] = iran_price['Population'] * 1000
      del iran_price['Unit']
```

```
[15]: color_dict = {'Akbar Hashemi Rafsanjani': "#FF0000",
                    'Mohammad Khatami': "#008000",
                    'Mahmoud Ahmadinejad': "#FF8C00",
                    'Hassan Rouhani': "#9400D3"}
```

```
[16]: list_of_items = list(iran_price['Item'].unique())
      for item in list_of_items:
          df_temp = iran_price[iran_price['Item'] == item]
          sns.set(style="whitegrid")
          plt.figure(figsize=(13, 5))
          sns.pointplot(x = 'Year' , y = 'Price per Kg in Rial' , hue = 'President' ,␣
      ↪palette= color_dict , data = df_temp)
          plt.ylabel('Value in Rial (IRR)')
          plt.xlabel('Year')
          plt.xticks(rotation = 90)
          plt.title(item)
          plt.show()
```
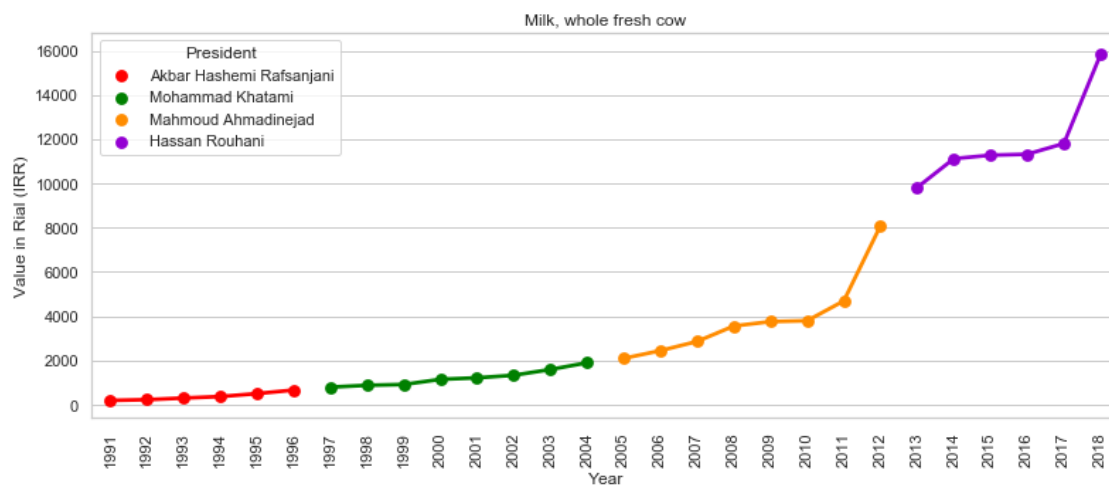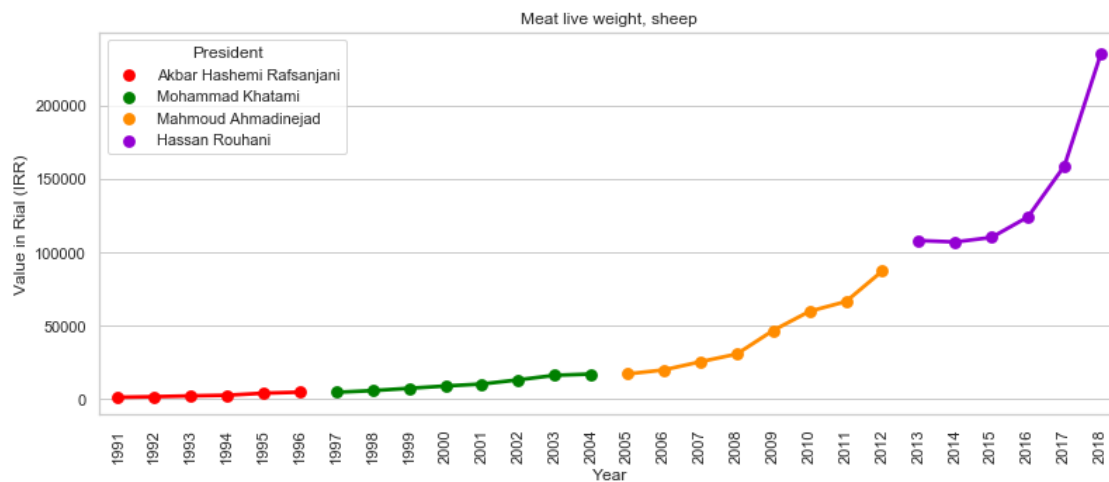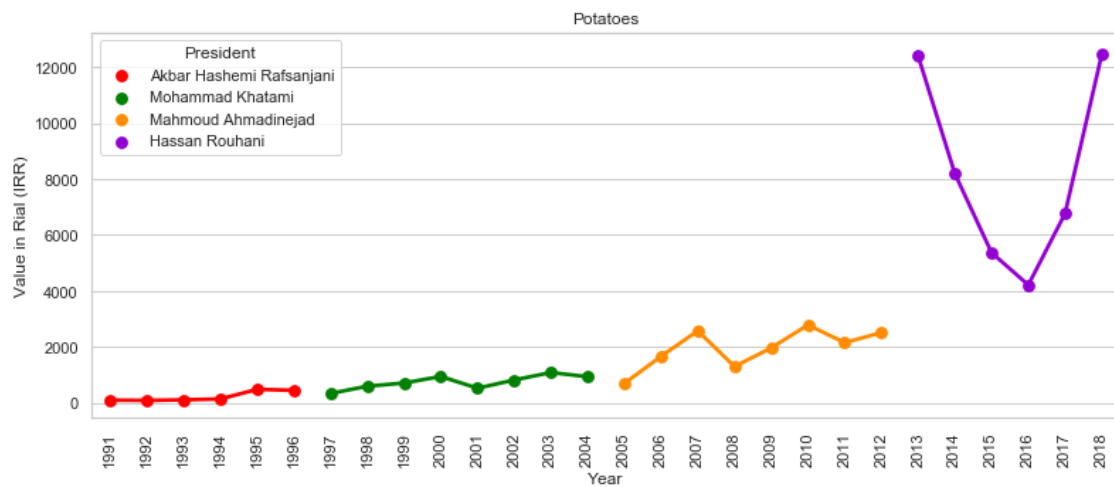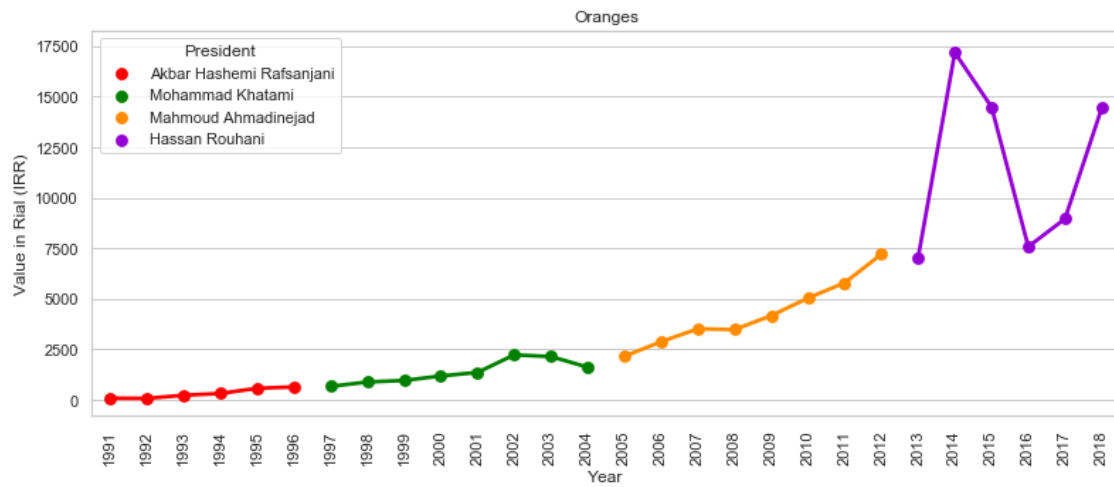
Beans, dry

Dates

Eggplants (aubergines)

Fruit, citrus nes

Fruit, fresh nes

Garlic

Honey, natural

Lemons and limes

Lentils

Meat live weight, cattle

Meat live weight, sheep

Milk, whole fresh cow

Onions, dry

Oranges

Potatoes

Rice, paddy

Soybeans

Sunflower seed

Tea

Tobacco, unmanufactured

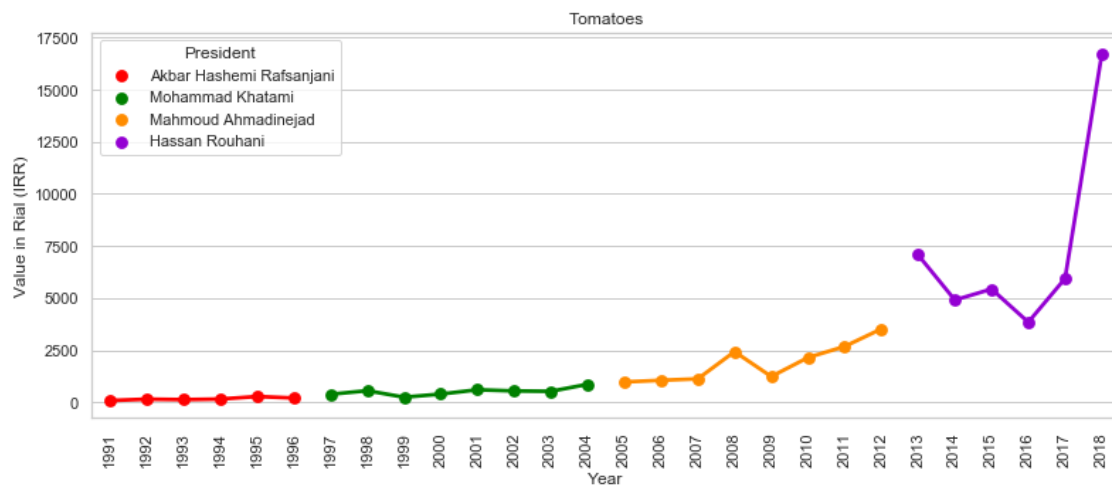Tomatoes

Vegetables, leguminous nes

Watermelons

Wheat
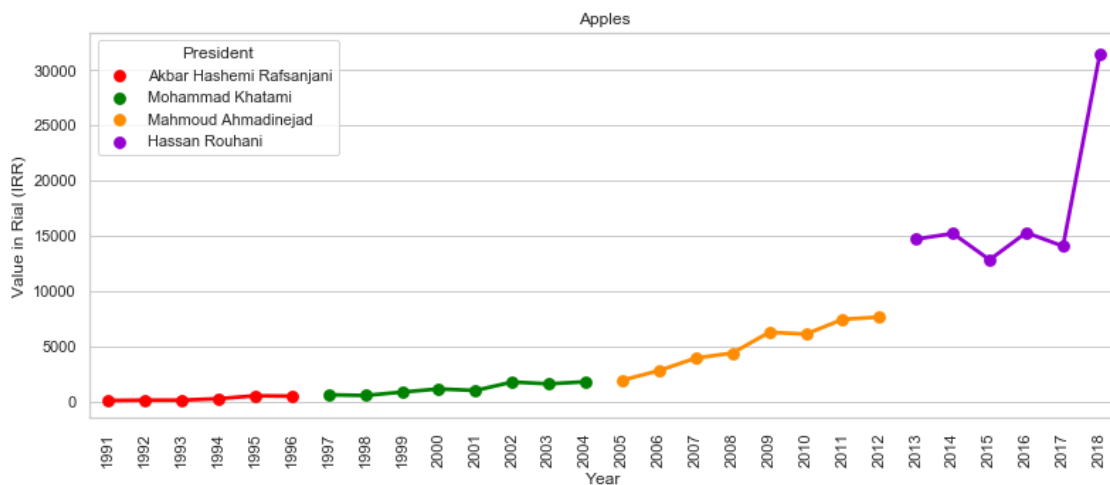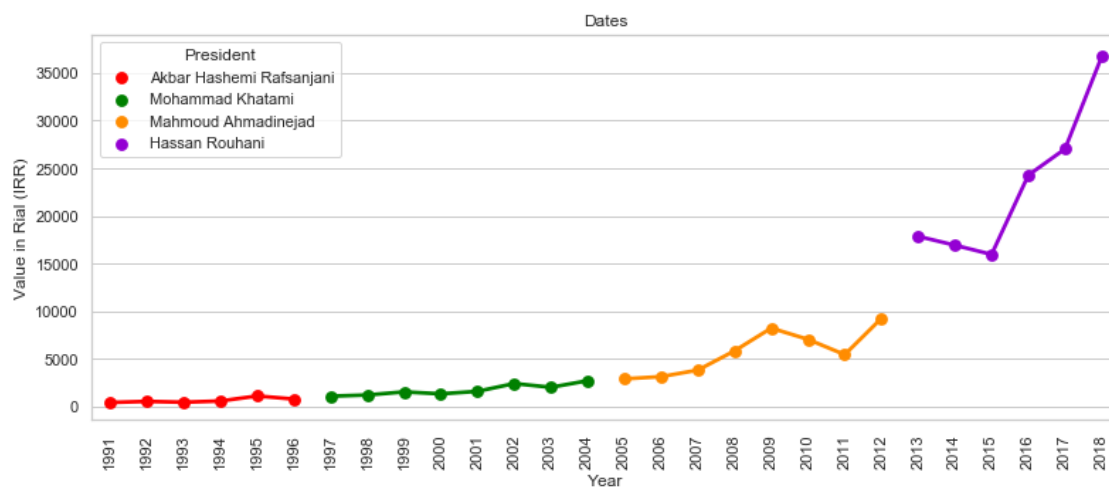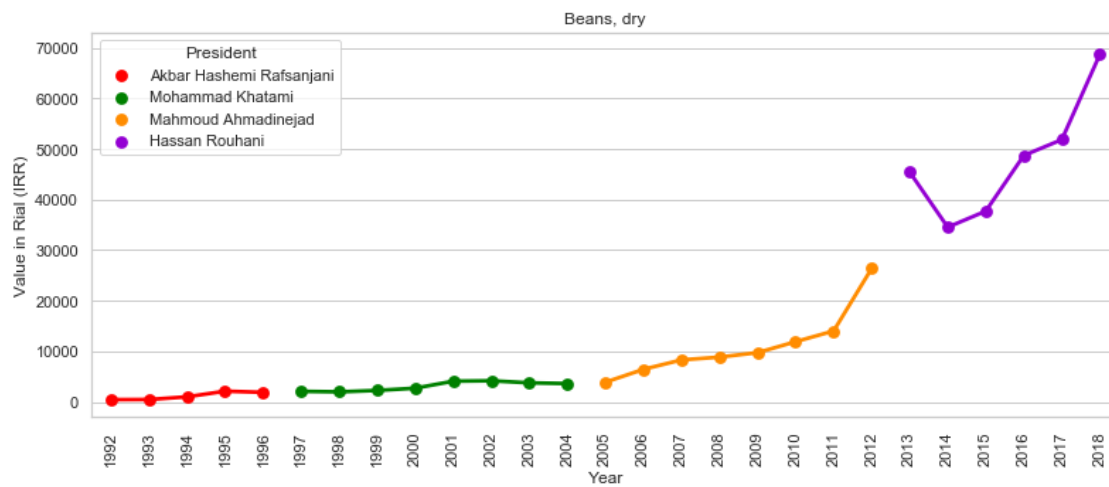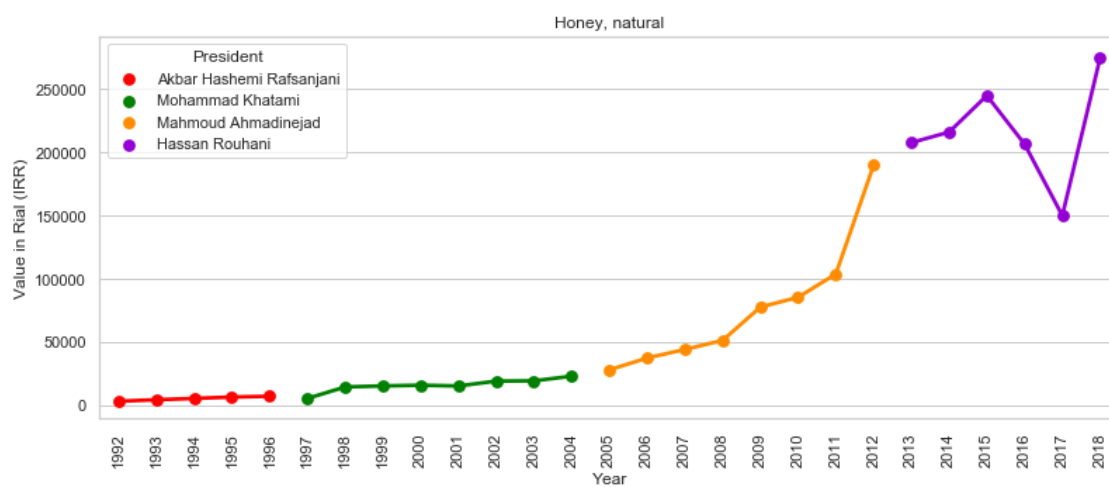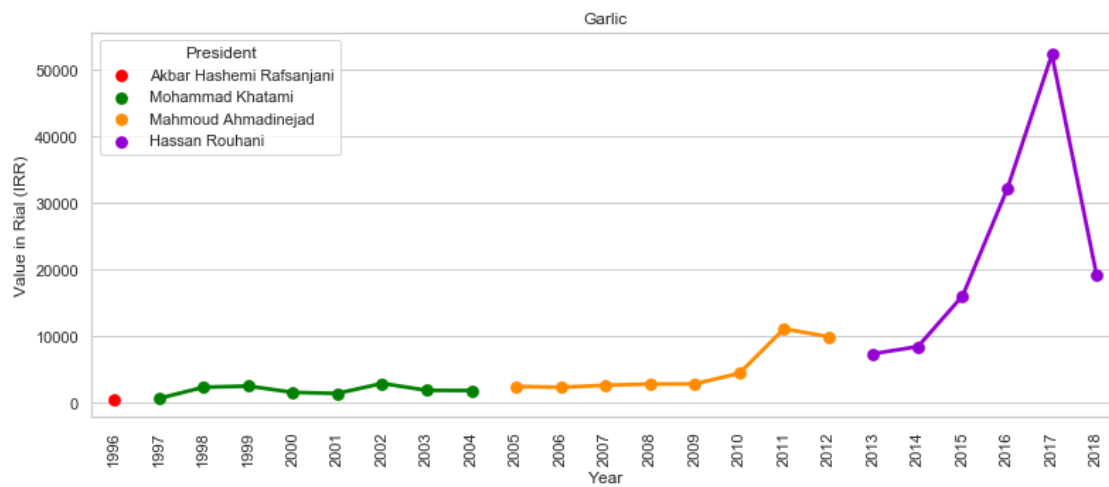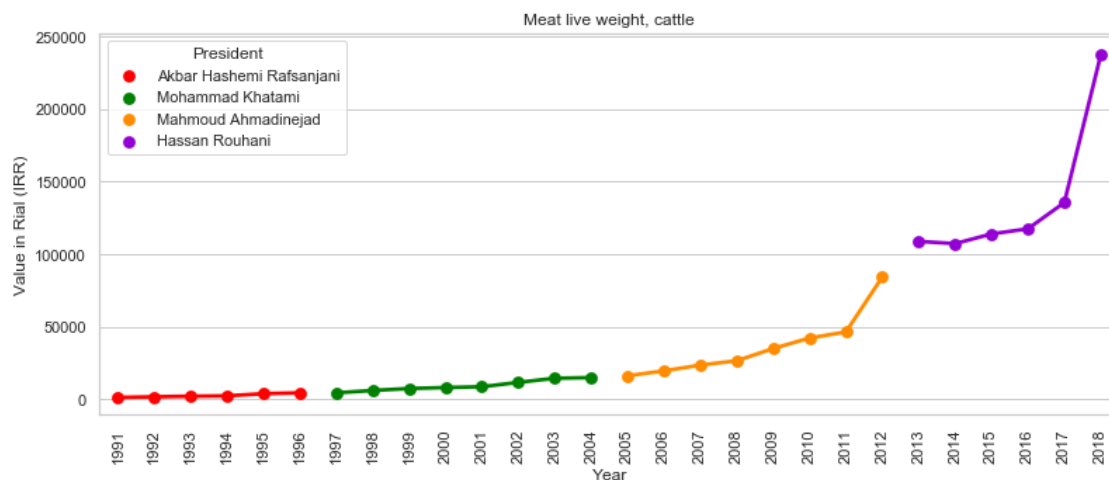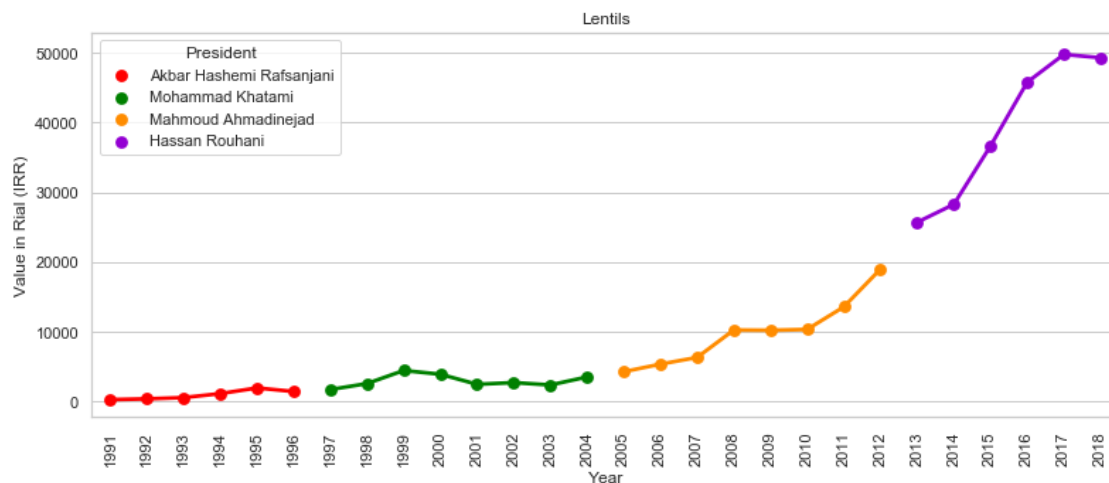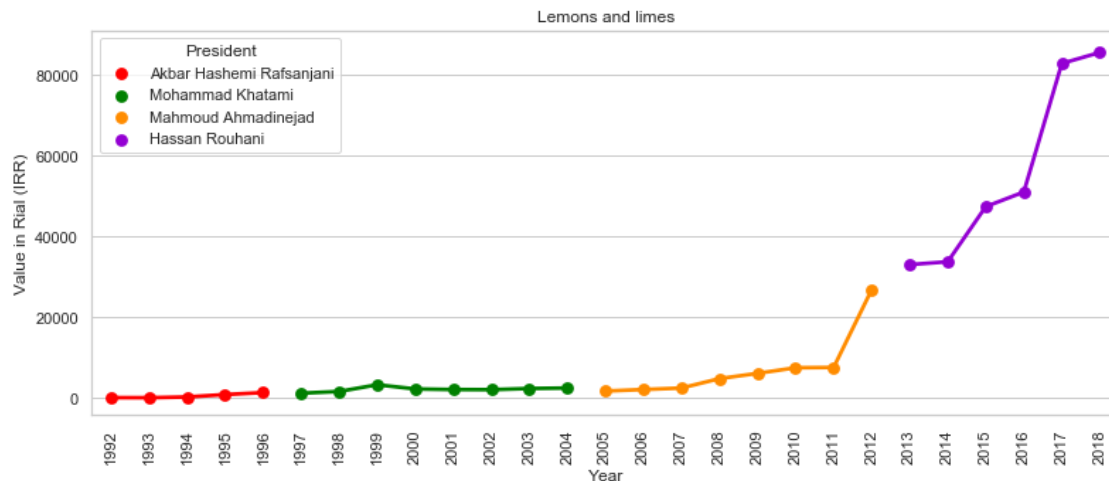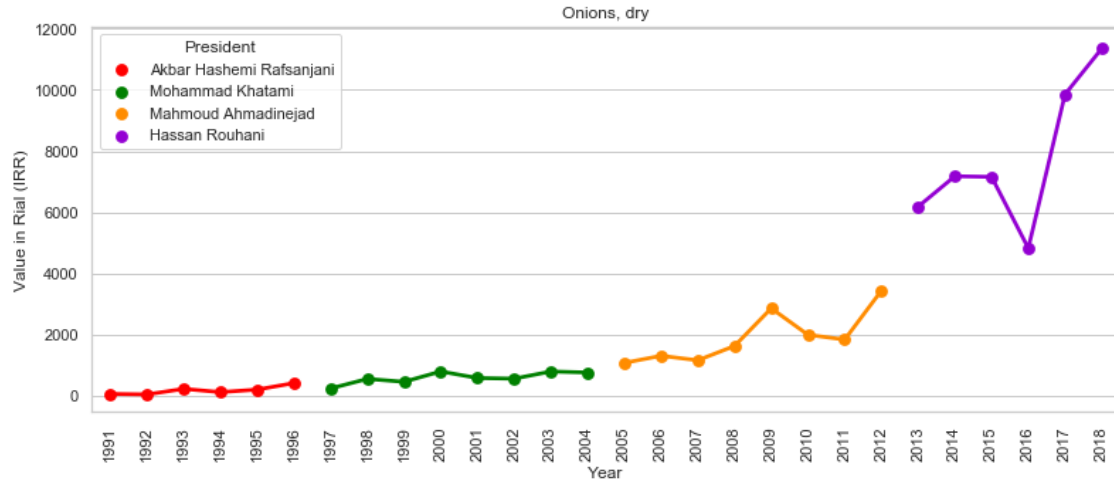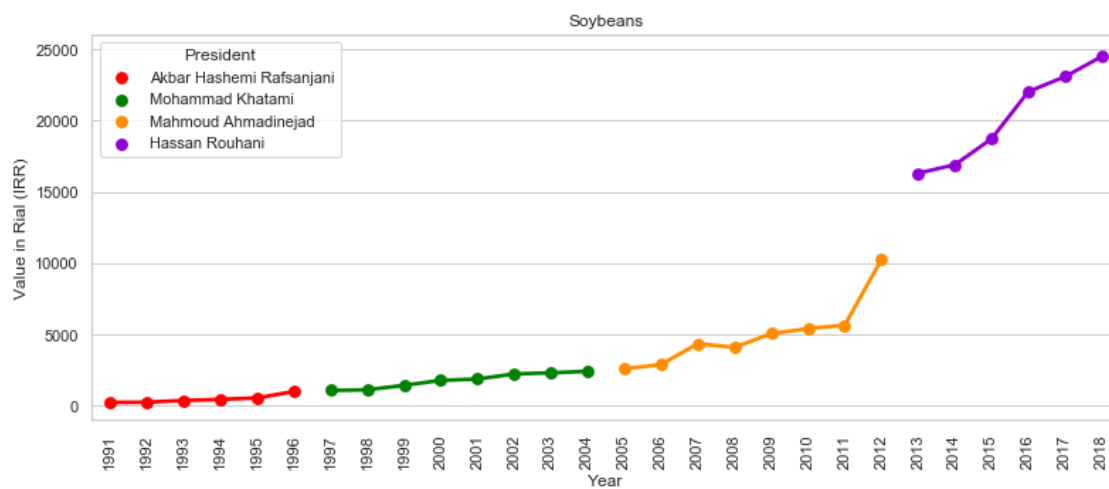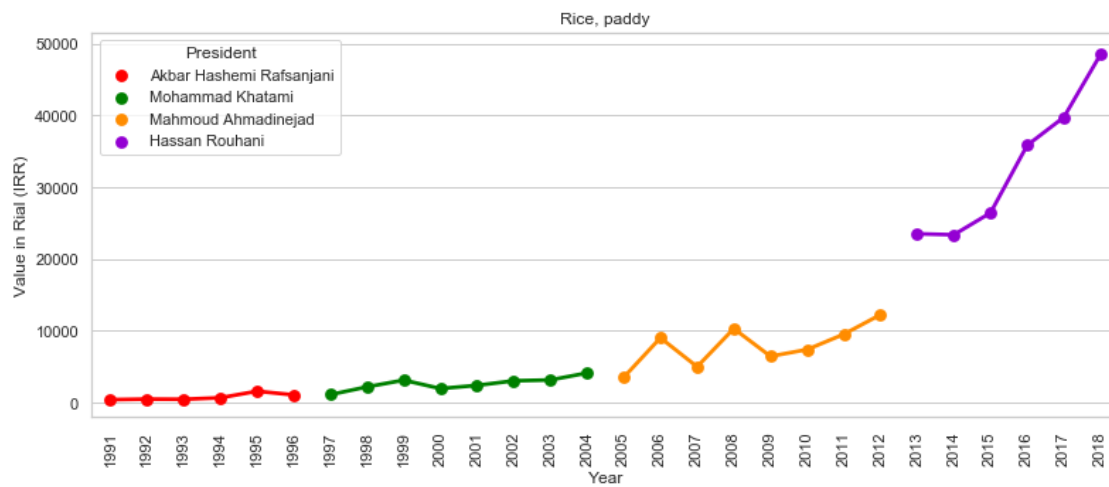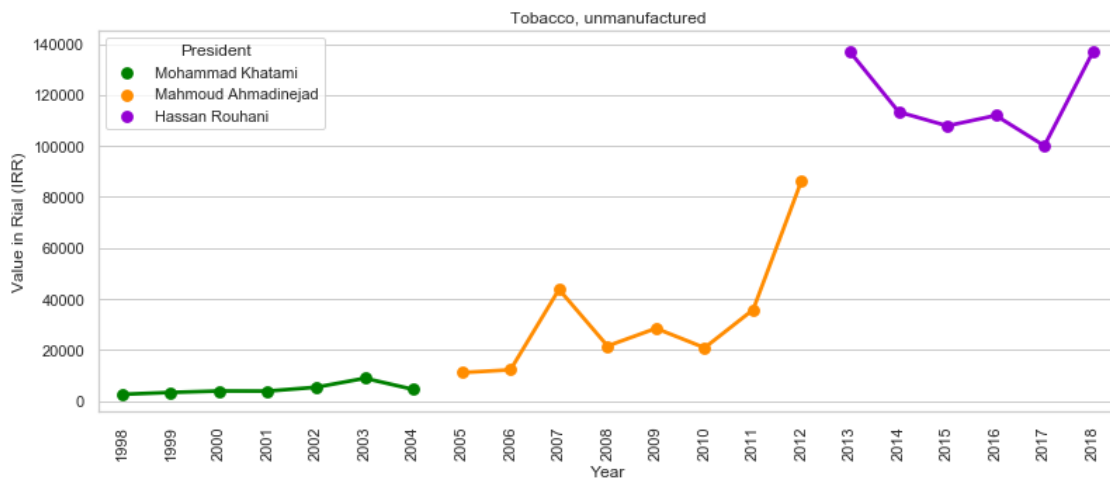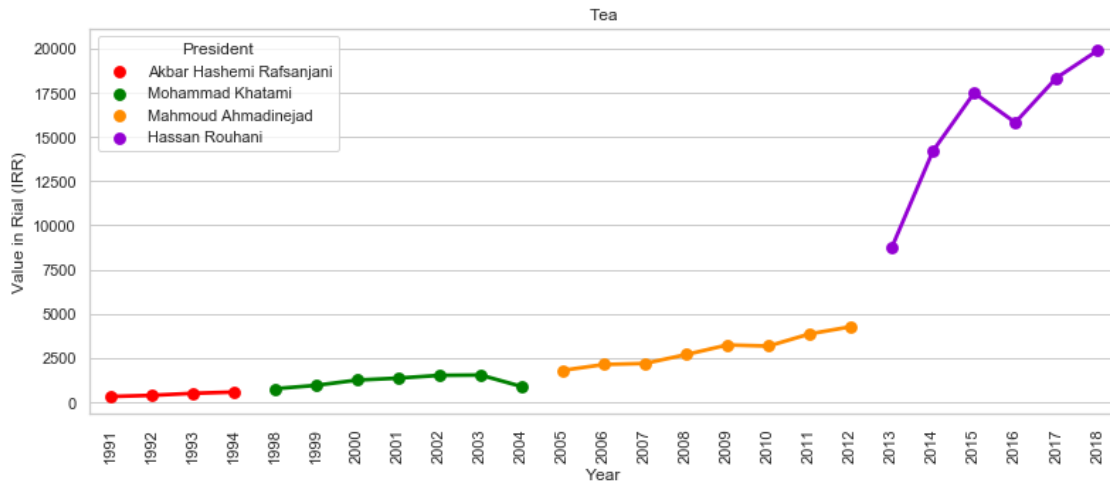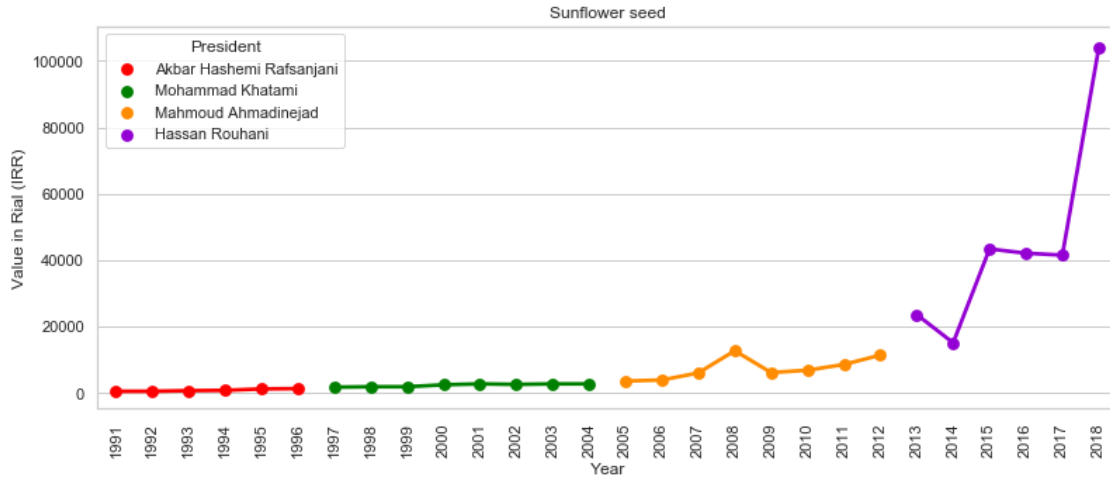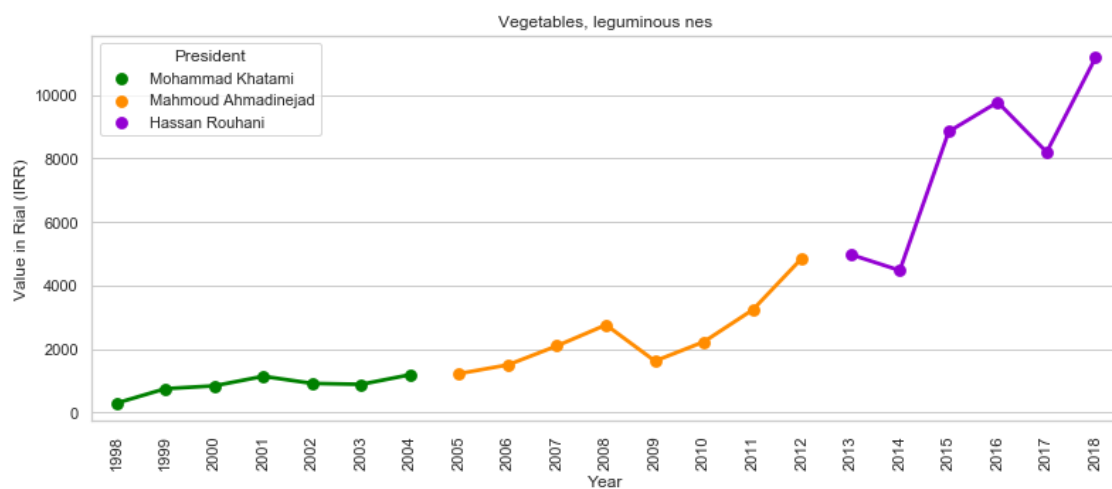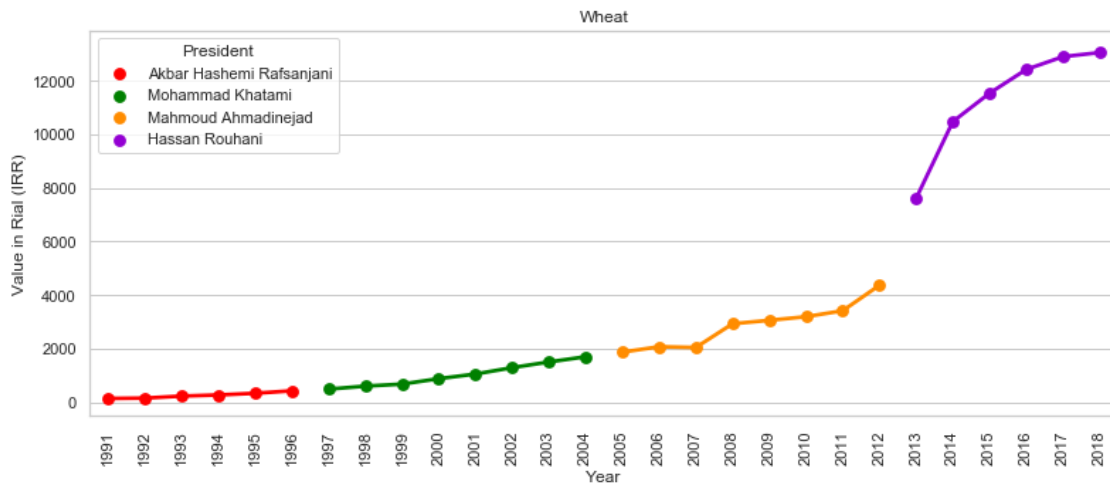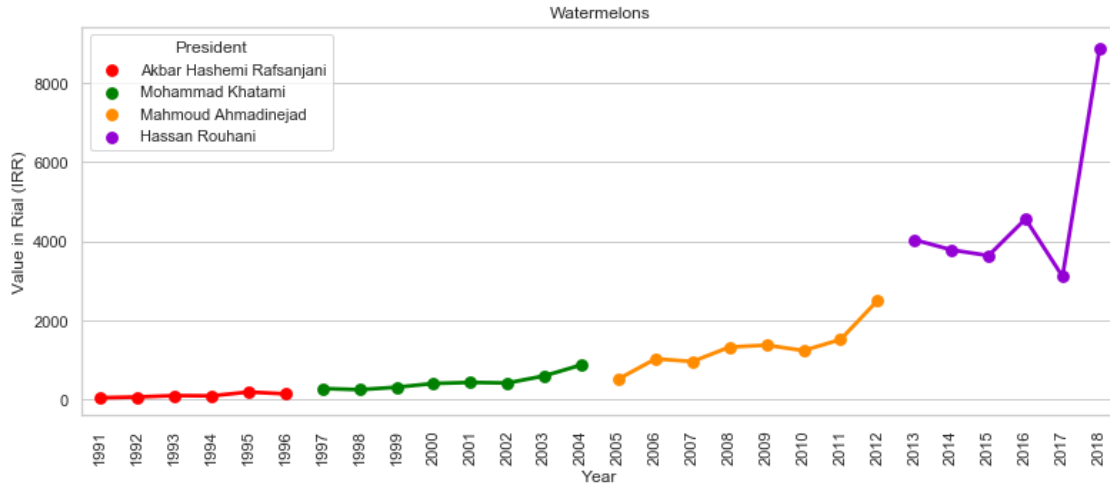

```
[17]: iran_exchange['President'] = 0
      for i in range(len(iran_exchange)):
          if iran_exchange.loc[i ,'Year'] >= 1991 and iran_exchange['Year'].iloc[i] <␣
      ↪1997:
              iran_exchange.loc[i, 'President'] = 'Akbar Hashemi Rafsanjani'
          elif iran_exchange.loc[i, 'Year'] >= 1997 and iran_exchange['Year'].iloc[i]␣
      ↪< 2005:
              iran_exchange.loc[i, 'President'] = 'Mohammad Khatami'
          elif iran_exchange.loc[i, 'Year'] >= 2005 and iran_exchange['Year'].iloc[i]␣
      ↪< 2013:
              iran_exchange.loc[i,'President'] = 'Mahmoud Ahmadinejad'
          elif iran_exchange.loc[i,'Year'] >= 2013 and iran_exchange['Year'].iloc[i]␣
      ↪<= 2020:
```

```
        iran_exchange.loc[i, 'President'] = 'Hassan Rouhani'
```

[18]:
```
iran_exchange = iran_exchange.loc[21:]
```

[19]:
```
sns.set(style="whitegrid")
plt.figure(figsize=(13, 5))
sns.pointplot(x = 'Year' , y = 'Value' , hue = 'President' , palette=␣
 ↪color_dict , data = iran_exchange)
plt.ylabel('Value of Dollar in Rial (IRR)')
plt.xlabel('Year')
plt.xticks(rotation = 90)
plt.title('Dollar Exchange')
plt.show()
```
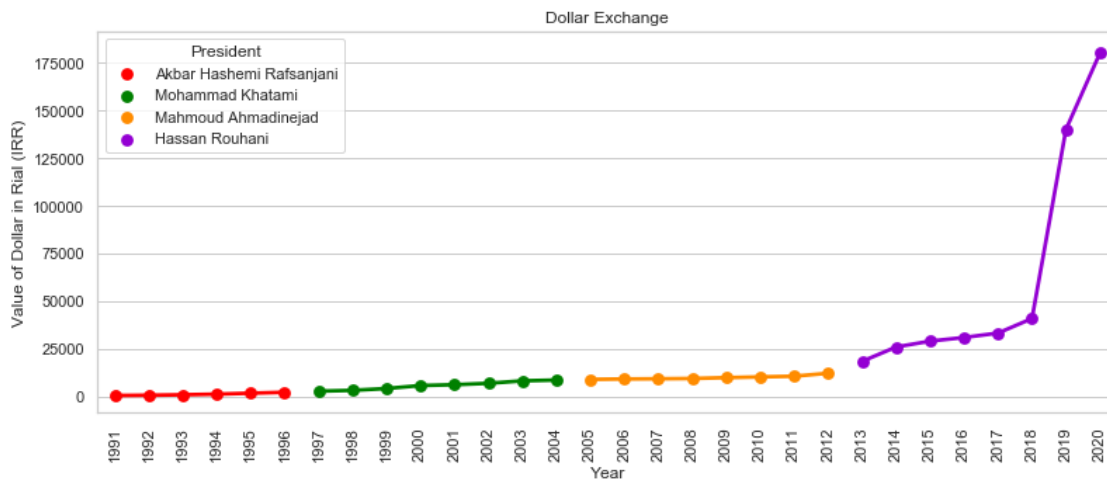

[20]:
```
iran_population.rename(columns={'Value': 'Population'}, inplace=True)
iran_population['Population'] = iran_population['Population'] * 1000
new_pop = {'Year': [2019,2020] , 'Population': [82913906 , 83992949] }
new_pop = pd.DataFrame(new_pop)
iran_population = iran_population.append(new_pop)
iran_population.reset_index(inplace=True)
del iran_population['index']
```

[21]:
```
iran_population['President'] = 0
for i in range(len(iran_population)):
    if iran_population.loc[i, 'Year'] >= 1991 and iran_population['Year'].
 ↪iloc[i] < 1997:
        iran_population.loc[i, 'President'] = 'Akbar Hashemi Rafsanjani'
    elif iran_population.loc[i, 'Year'] >= 1997 and iran_population['Year'].
 ↪iloc[i] < 2005:
        iran_population.loc[i, 'President'] = 'Mohammad Khatami'
```

```
    elif iran_population.loc[i ,'Year'] >= 2005 and iran_population['Year'].
→iloc[i] < 2013:
        iran_population.loc[i, 'President'] = 'Mahmoud Ahmadinejad'
    elif iran_population.loc[i, 'Year'] >= 2013 and iran_population['Year'].
→iloc[i] <= 2020:
        iran_population.loc[i, 'President'] = 'Hassan Rouhani'
```

```python
[22]: sns.set(style="whitegrid")
plt.figure(figsize=(13, 5))
sns.pointplot(x = 'Year' , y = 'Population' , hue = 'President' , palette=␣
→color_dict , data = iran_population)
plt.ylabel('Population')
plt.xlabel('Year')
plt.xticks(rotation = 90)
plt.title('Population of Iran')
plt.show()
```
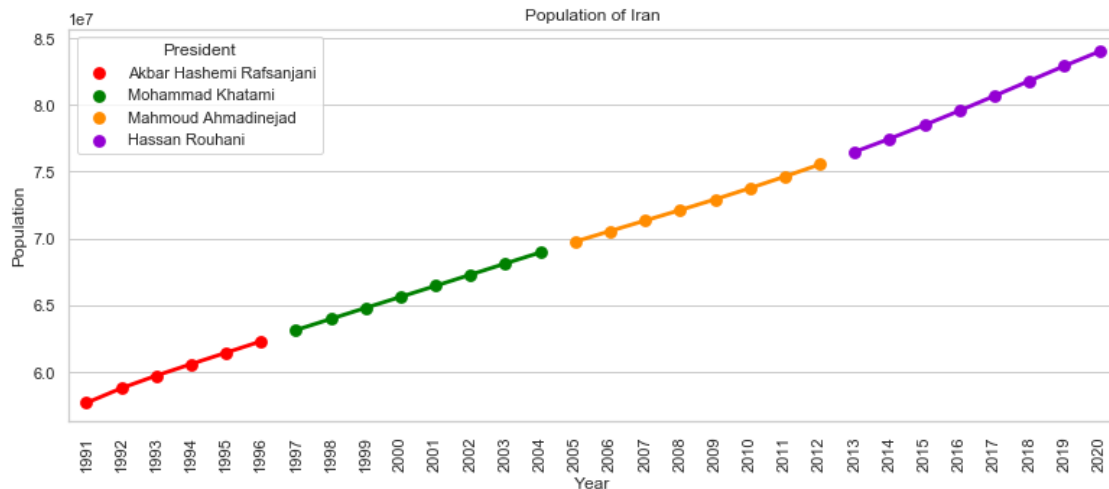


```python
[23]: iran_population["Population change"] = iran_population["Population"].
→pct_change()
```

```python
[24]: plt.figure(figsize=(13, 5))
sns.pointplot(x= "Year", y= "Population change", data= iran_population)
plt.xlabel("Year")
plt.xticks(rotation= 90)
plt.ylabel("Population changes")
plt.title("Population change by Year")
plt.show()
```
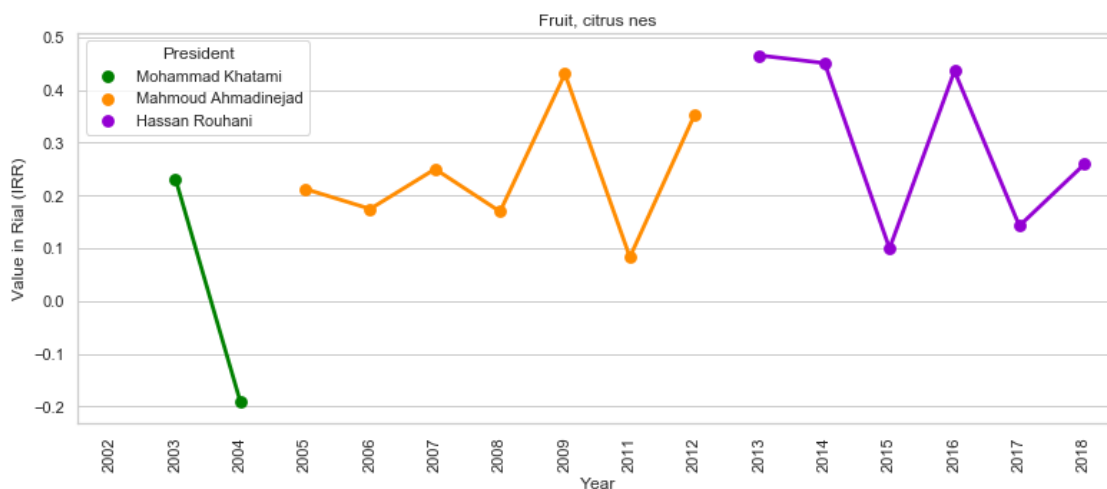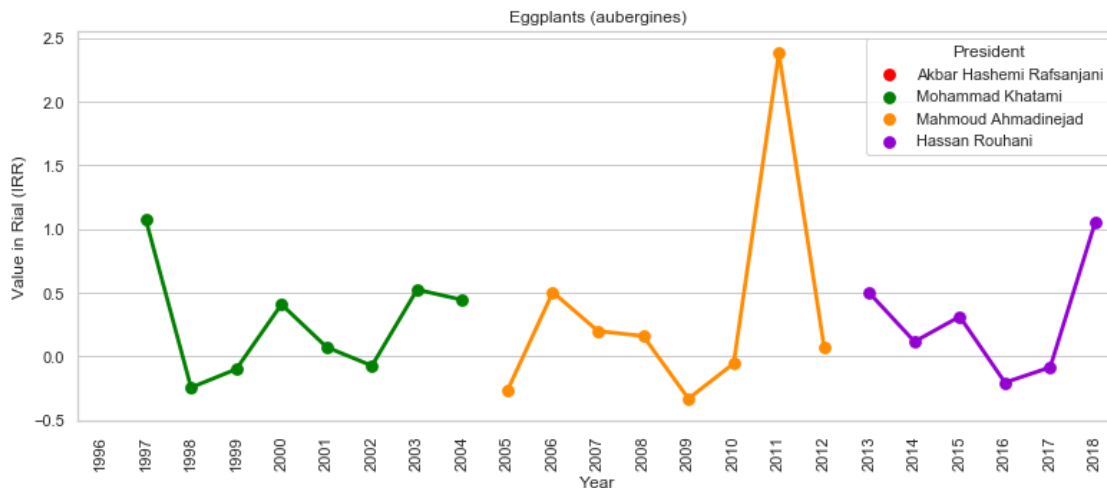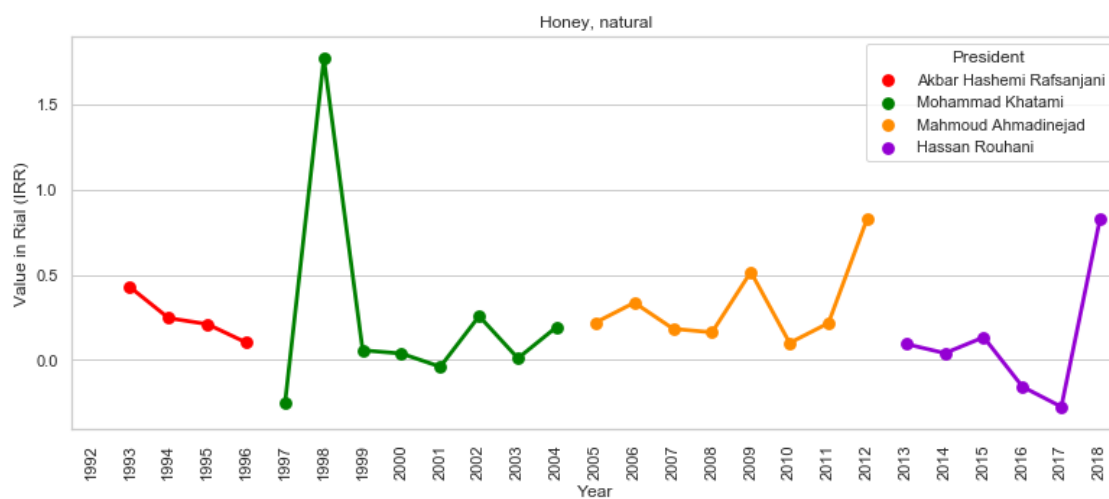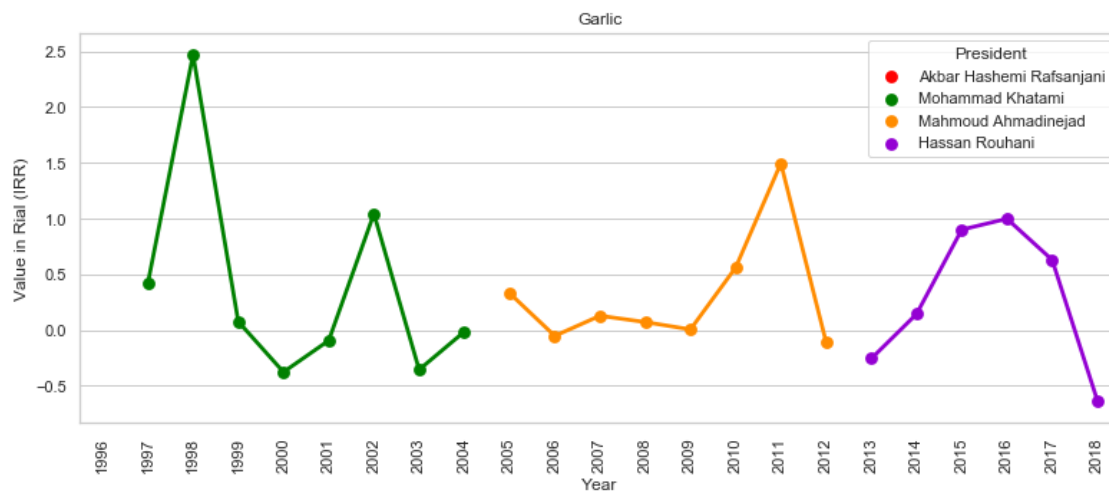
16

Population change by Year
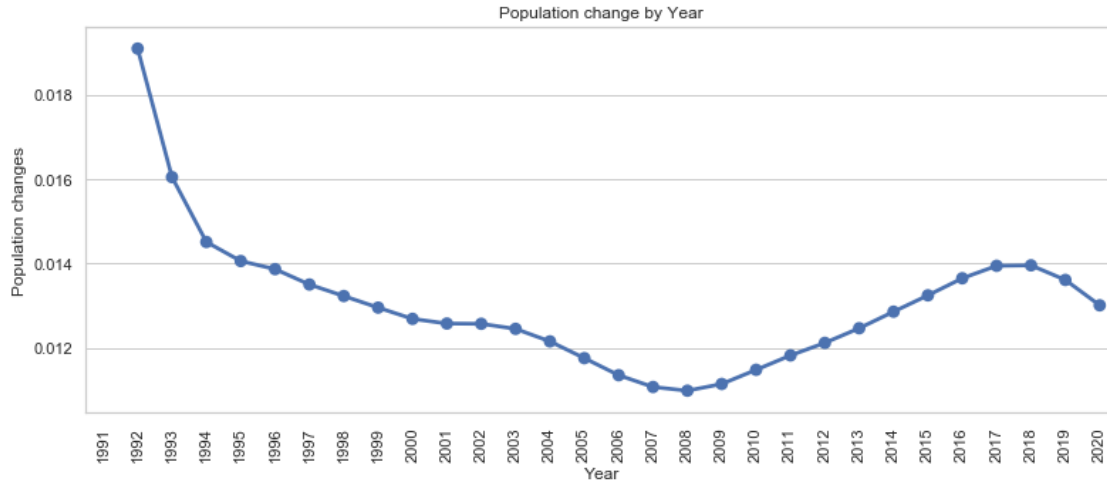
```
[25]: iran_price["Item change"] = iran_price.groupby("Item")["Price per Kg in Rial"].
      ↪pct_change()
```
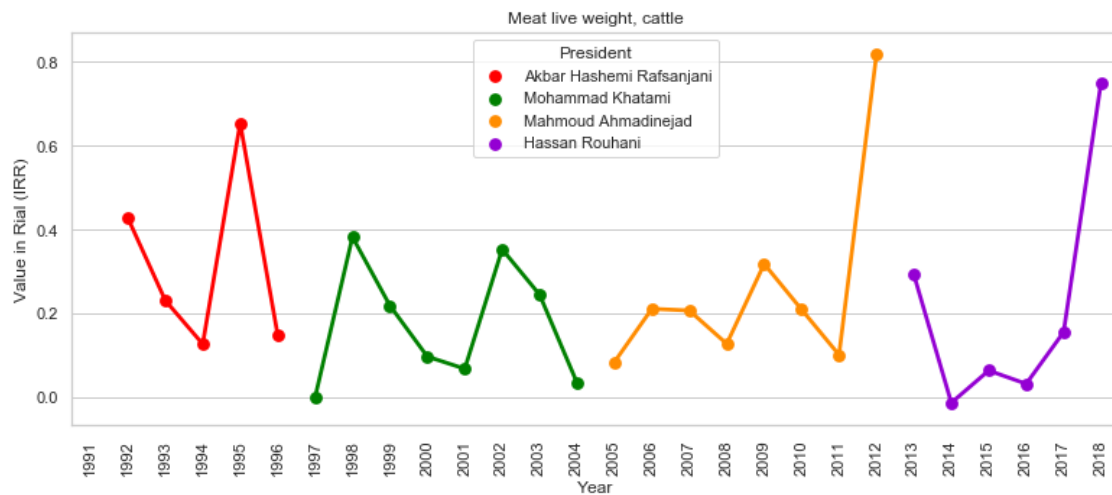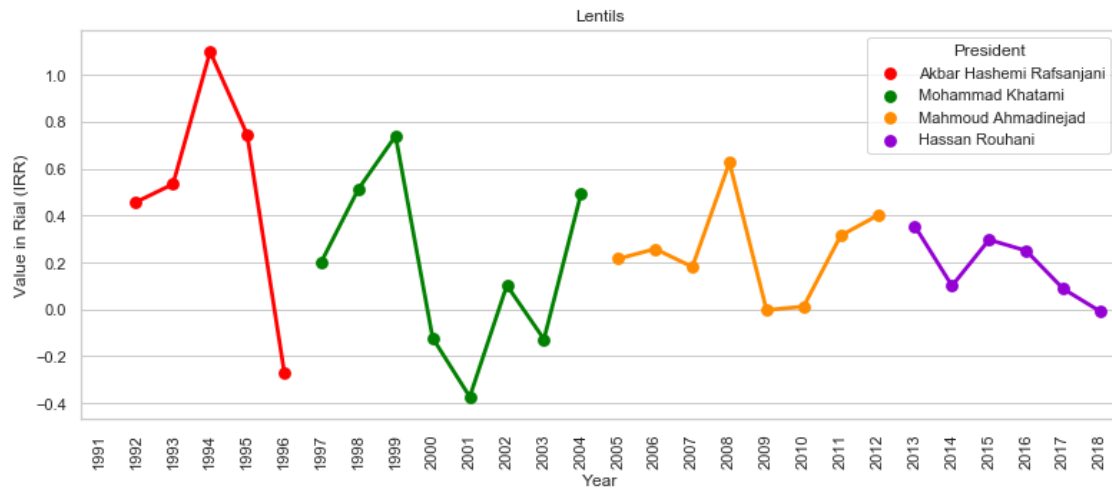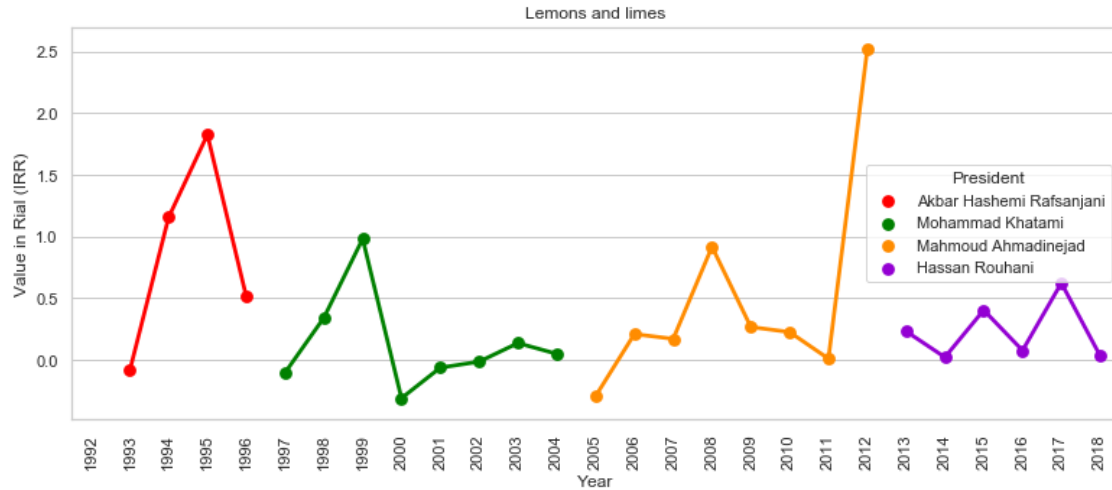
```
[26]: for item in list_of_items:
          df_temp = iran_price[iran_price['Item'] == item]
          sns.set(style="whitegrid")
          plt.figure(figsize=(13, 5))
          sns.pointplot(x = 'Year' , y = "Item change" , hue = 'President' , palette=
      ↪color_dict , data = df_temp)
          plt.ylabel('Value in Rial (IRR)')
          plt.xlabel('Year')
          plt.xticks(rotation = 90)
          plt.title(item)
          plt.show()
```
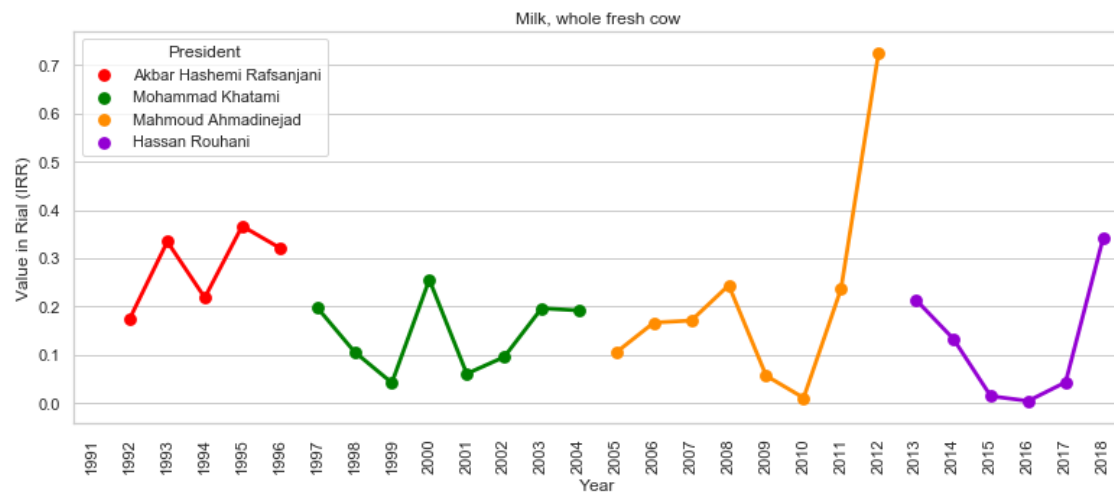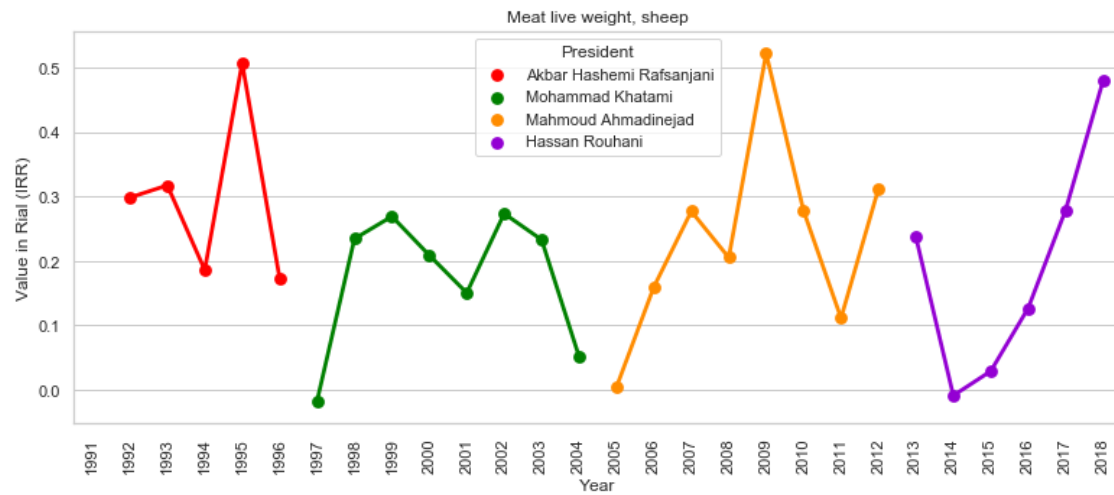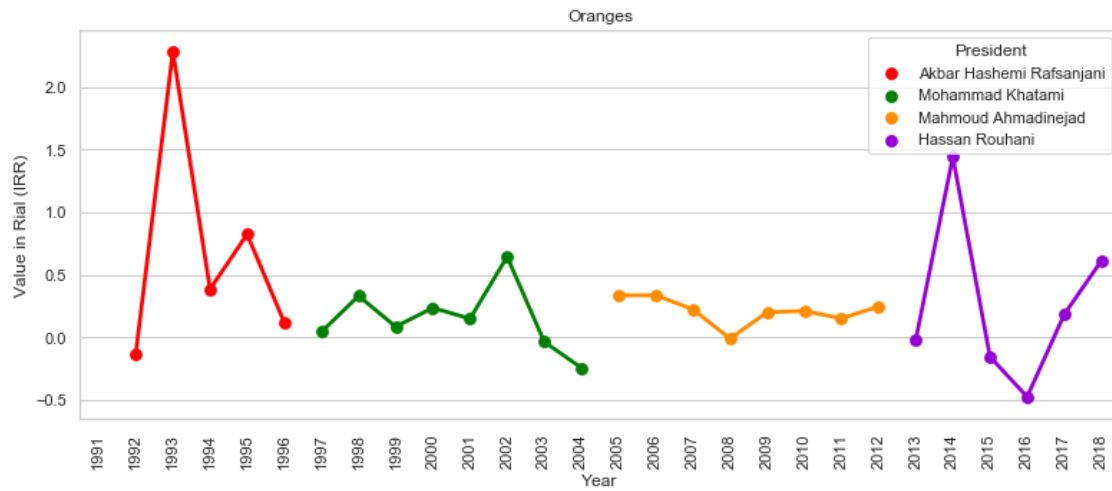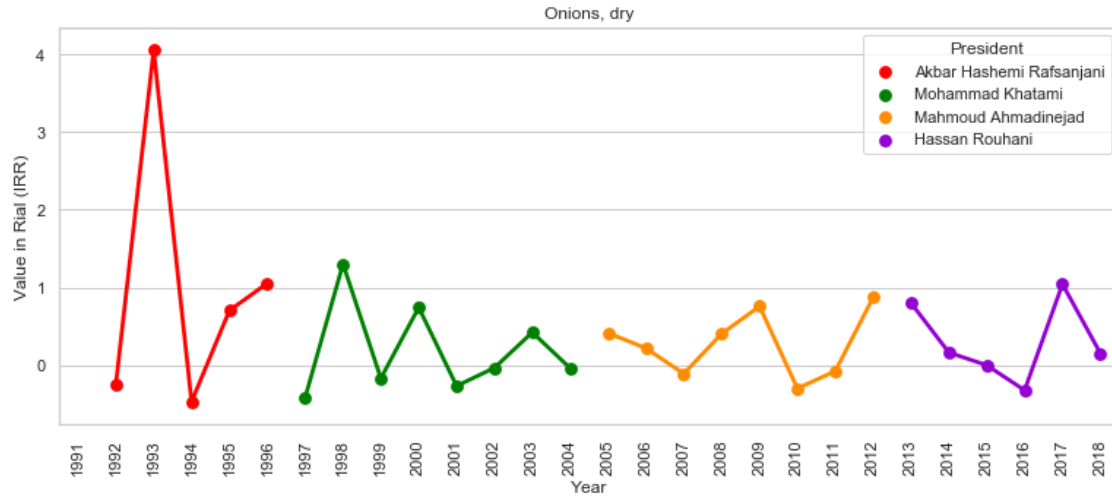

Apples

Beans, dry

Dates

Eggplants (aubergines)

Fruit, citrus nes

Fruit, fresh nes

Garlic

Honey, natural

Lemons and limes

Lentils

Meat live weight, cattle

Meat live weight, sheep

Milk, whole fresh cow

Onions, dry

Oranges

Potatoes

Rice, paddy

Soybeans

Sunflower seed

Tea

Tobacco, unmanufactured

Tomatoes

Vegetables, leguminous nes

Watermelons

Wheat
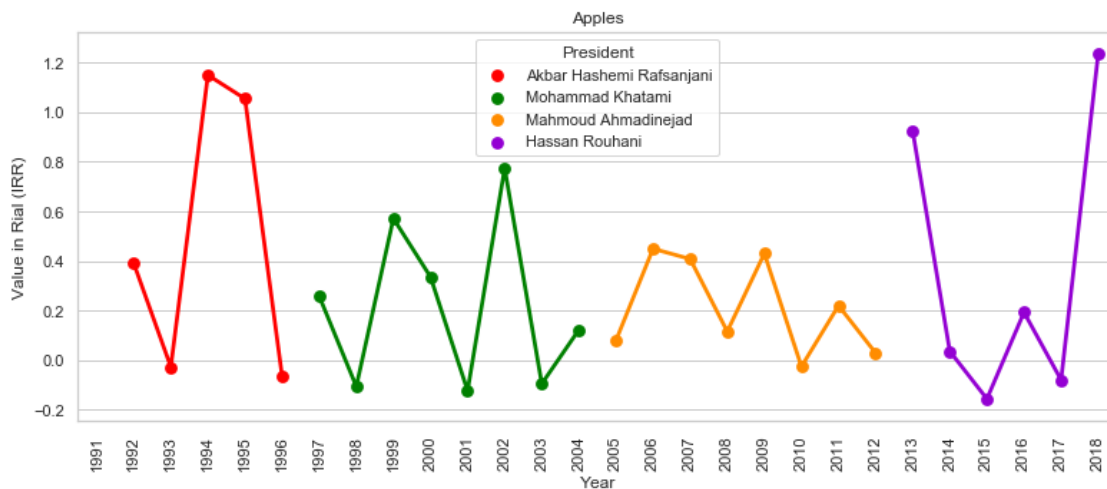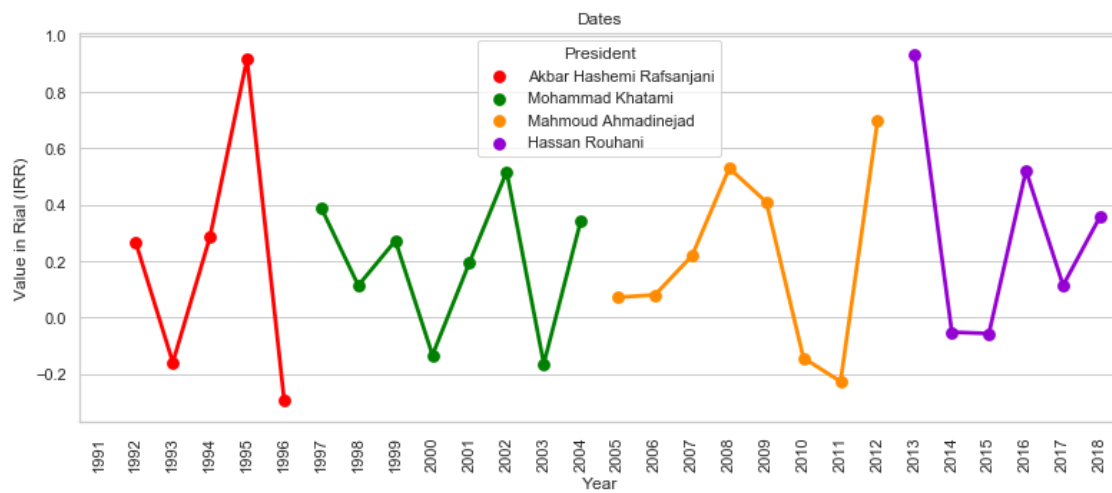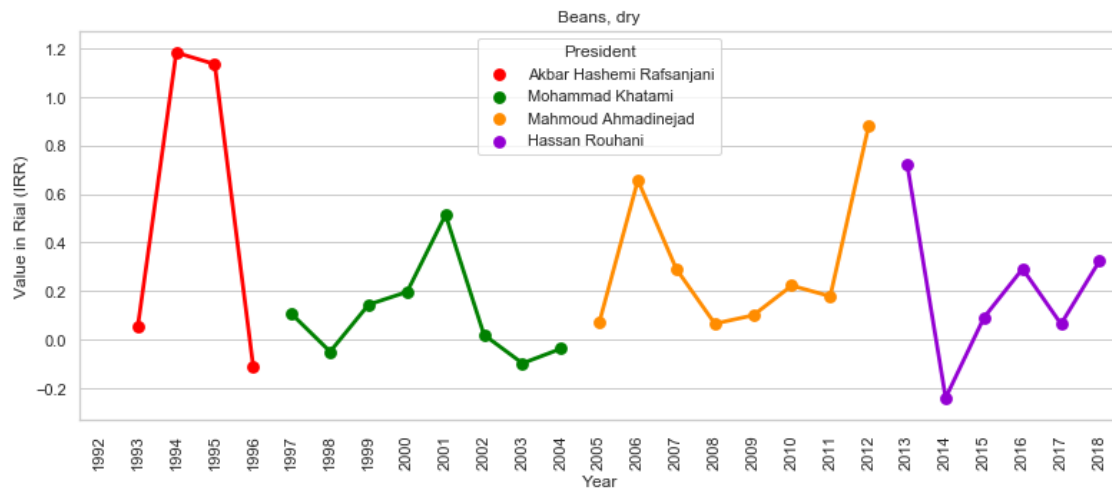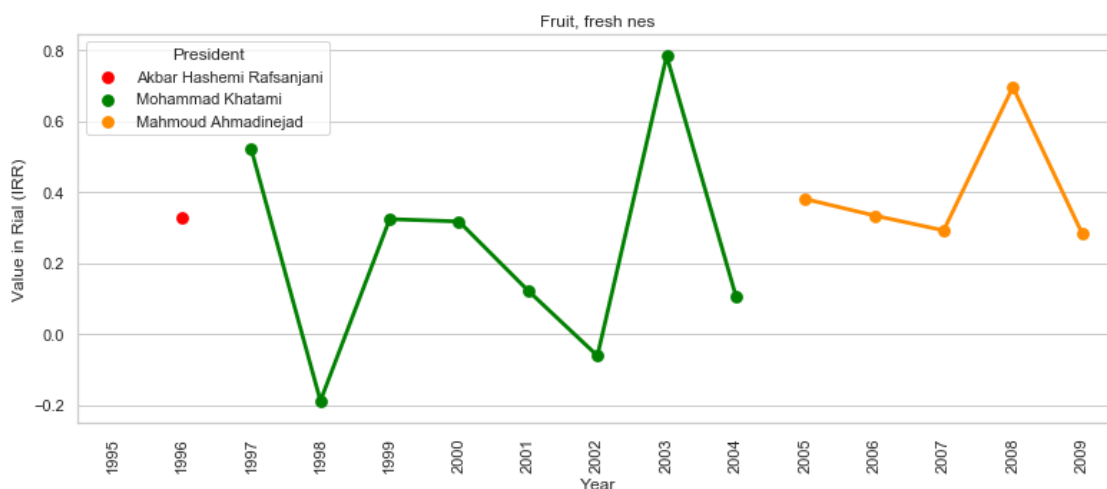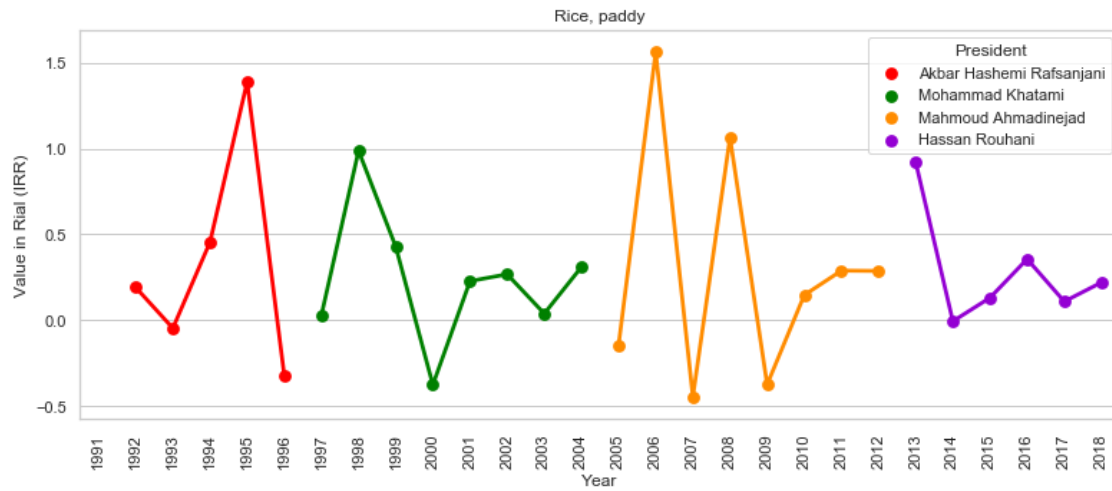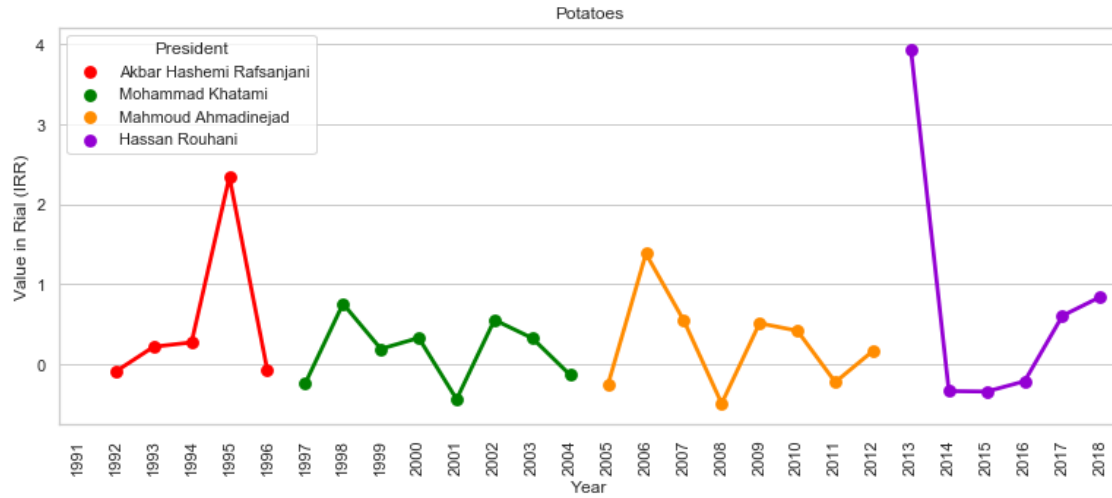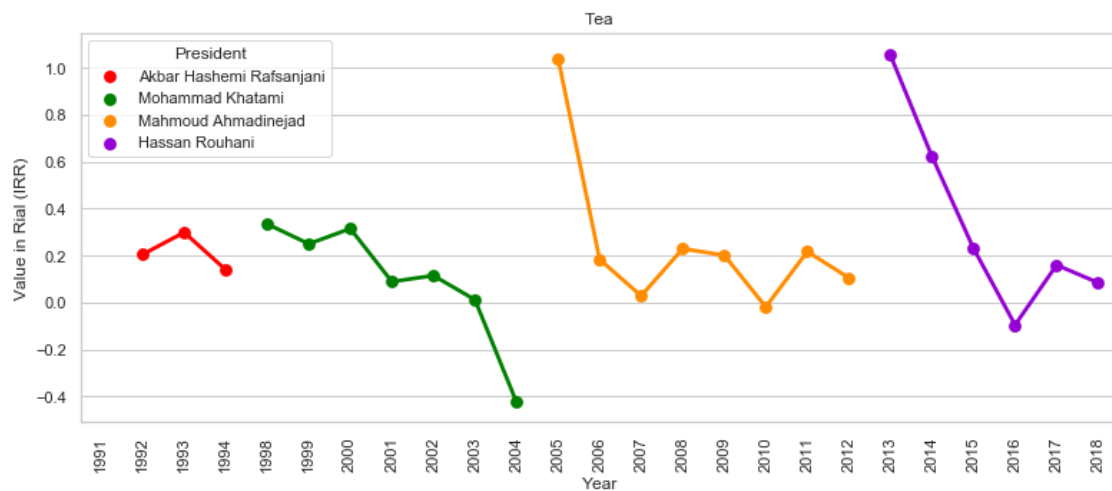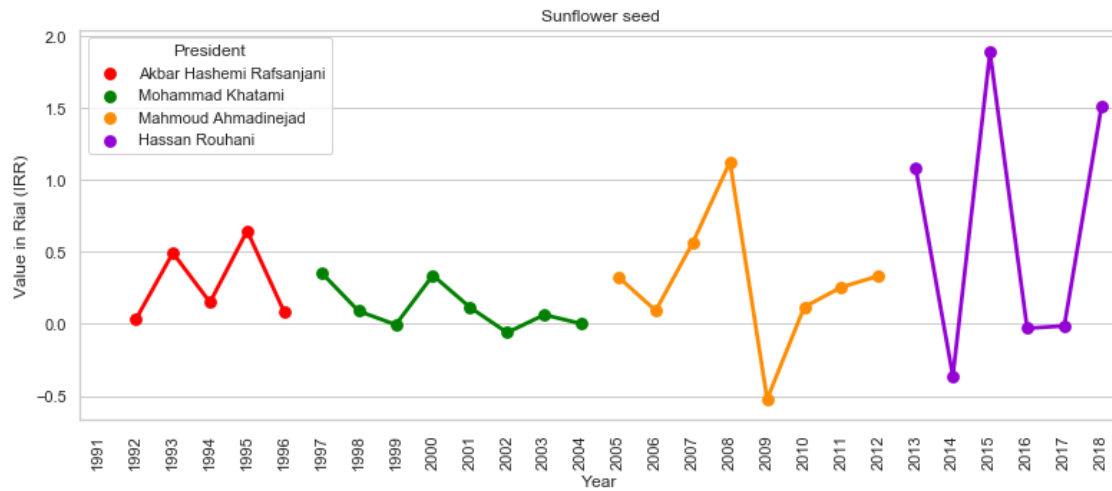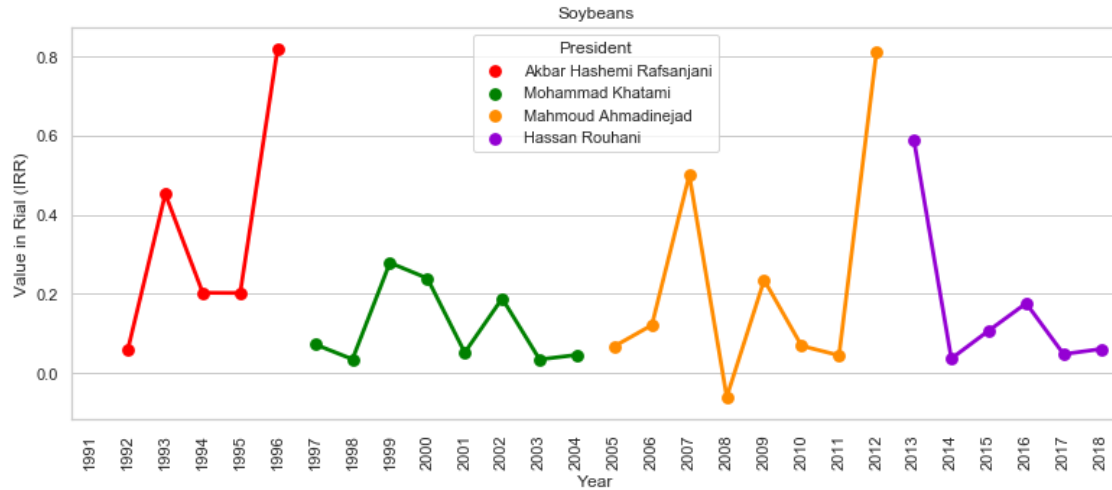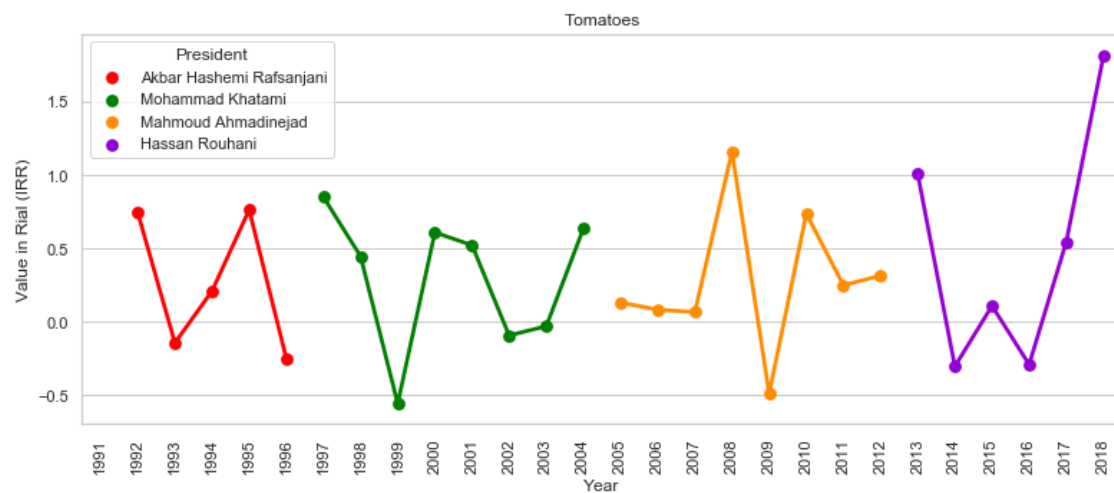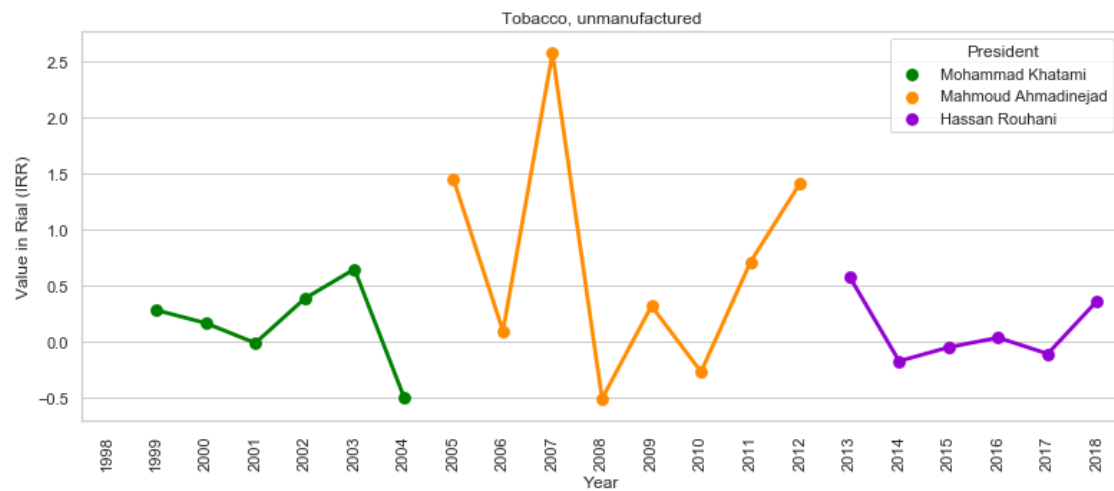
```
[27]: home_basket_per_year = iran_price.groupby("Year")["Price per Kg in Rial"].sum().
      ↪to_frame().reset_index()
```
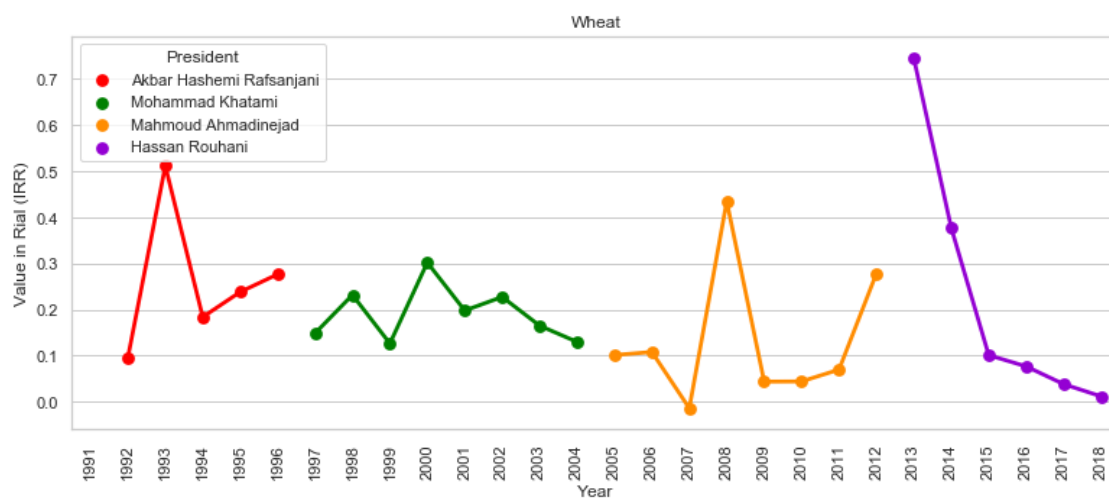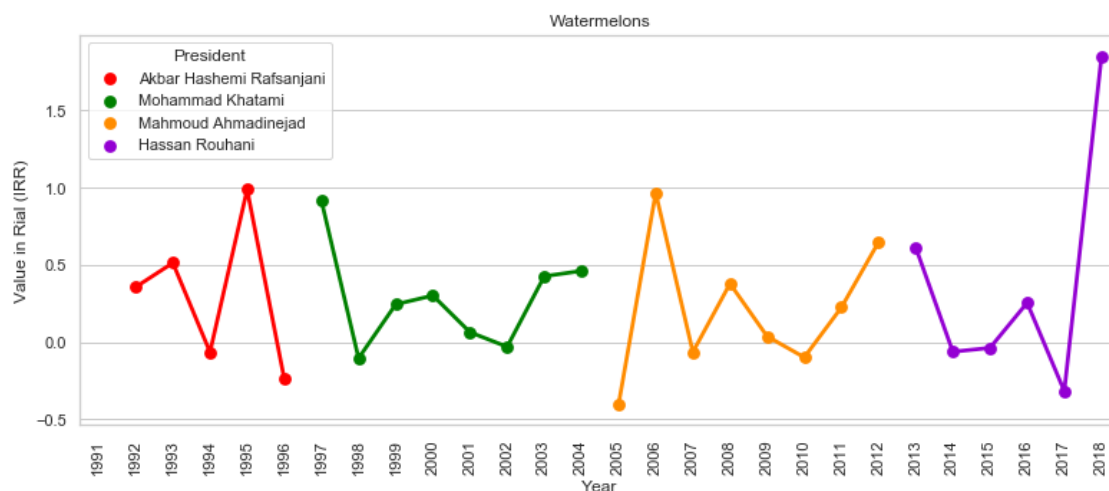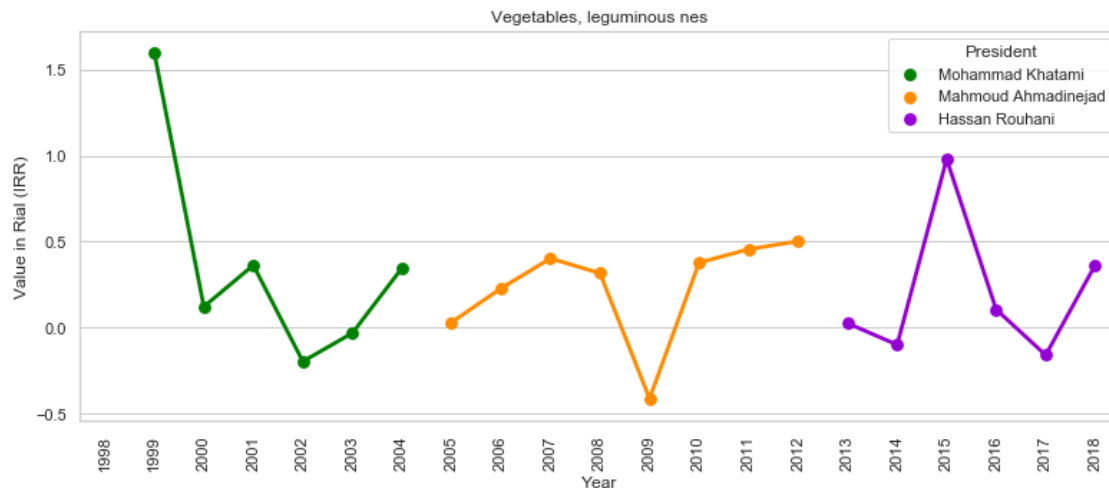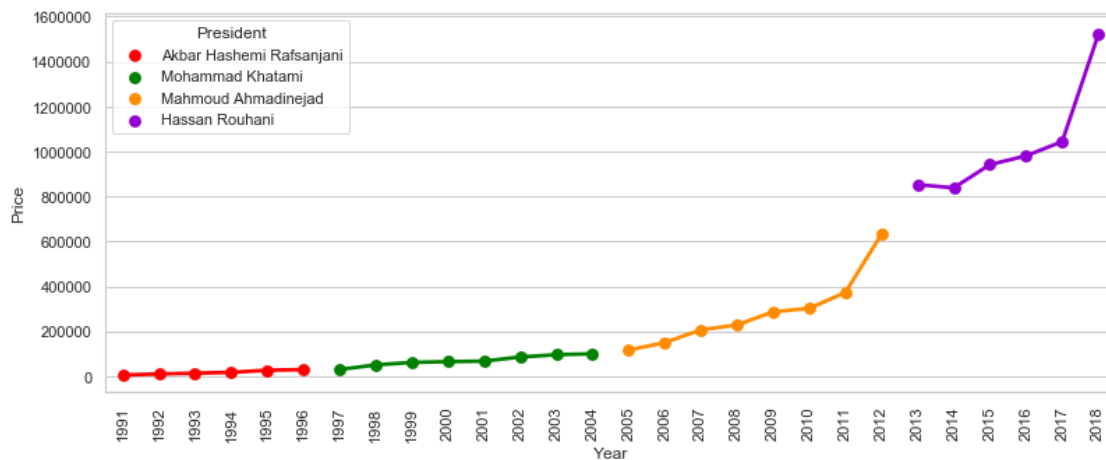
```
[28]: home_basket_per_year['President'] = 0
      for i in range(len(home_basket_per_year)):
          if home_basket_per_year.loc[i, 'Year'] >= 1991 and␣
      ↪home_basket_per_year['Year'].iloc[i] < 1997:
              home_basket_per_year.loc[i, 'President'] = 'Akbar Hashemi Rafsanjani'
          elif iran_population.loc[i, 'Year'] >= 1997 and␣
      ↪home_basket_per_year['Year'].iloc[i] < 2005:
              home_basket_per_year.loc[i, 'President'] = 'Mohammad Khatami'
          elif iran_population.loc[i ,'Year'] >= 2005 and␣
      ↪home_basket_per_year['Year'].iloc[i] < 2013:
              home_basket_per_year.loc[i, 'President'] = 'Mahmoud Ahmadinejad'
          elif home_basket_per_year.loc[i, 'Year'] >= 2013 and␣
      ↪home_basket_per_year['Year'].iloc[i] <= 2020:
              home_basket_per_year.loc[i, 'President'] = 'Hassan Rouhani'
```

```
[29]: plt.figure(figsize=(13, 5))
      sns.pointplot(x= "Year", y= "Price per Kg in Rial", hue= "President" , palette=␣
      ↪color_dict , data= home_basket_per_year)
      plt.xticks(rotation= 90)
      plt.xlabel("Year")
      plt.ylabel("Price")
      plt.show()
```



```
[ ]:
```

[ ]: