МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «ВГУ»)

Факультет компьютерных наук

*Кафедра* *программирования и информационных технологий*

*Прогнозирование коротких временных рядов*

Отчет по лабораторной работе № *2*

*09.03.02 Информационные системы и технологии*

*Программная инженерия в информационных системах*

Отчёт составил:

Свиридов Фёдор Юрьевич, группа 5.2, вариант 11

Воронеж2023

**Задание:**

1. Построить график исходного временного ряда (а также построить прогноз потребления крепленых вин (fort) на 8 месяцев)

2. Ответить на следующие вопросы:

* Есть ли у ряда тренд?
* Есть ли сезонность? И какая она?
* Меняет ли ряд свой характер?
* Есть ли в данных выбросы?

3. Построить прогностическую модель.

4. Выполнить прогноз на заданный промежуток времени.

5. Построить графики исходного ряда и подогнанных данных с учетом прогноза.

6. Привести набор чисел — спрогнозированных значений рассматриваемой величины для каждого из будущих моментов времени.

Содержимое файла wine\_Austral.dat:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| fort | dry | sweet | red | rose | spark | total | year\_ | month\_ | date\_ |
| 2585 | 1954 | 85 | 464 | 112 | 1686 | 15136 | 1980 | 1 | JAN 1980 |
| 3368 | 2302 | 89 | 675 | 118 | 1591 | 16733 | 1980 | 2 | FEB 1980 |
| 3210 | 3054 | 109 | 703 | 129 | 2304 | 20016 | 1980 | 3 | MAR 1980 |
| 3111 | 2414 | 95 | 887 | 99 | 1712 | 17708 | 1980 | 4 | APR 1980 |

//и так далее

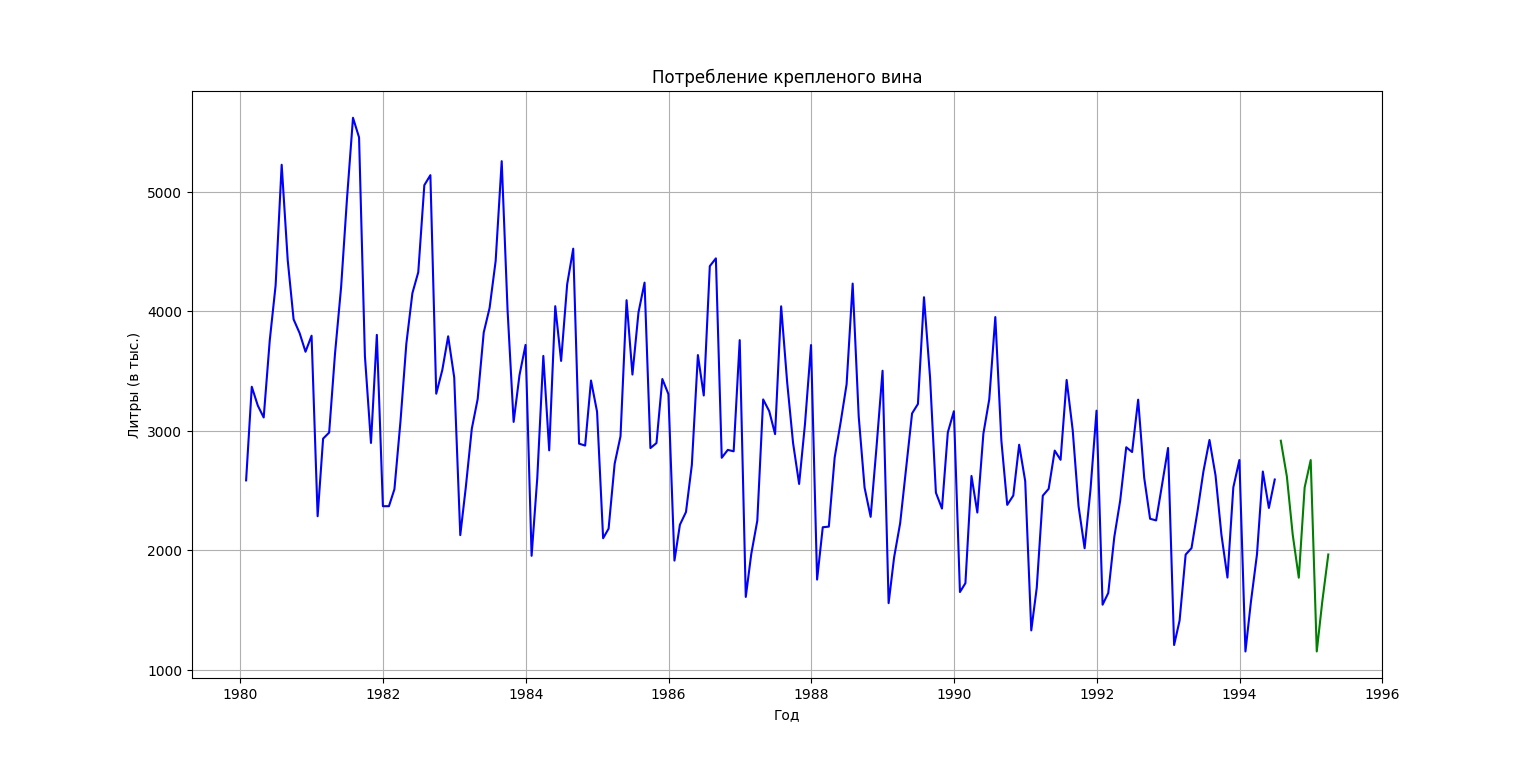
**Код проекта на Python:**

import warnings  
from datetime import datetime  
from calendar import isleap  
from statsmodels.tsa.seasonal import seasonal\_decompose  
from statsmodels.tsa.statespace.sarimax import SARIMAX  
  
import pandas as pd  
import matplotlib  
import matplotlib.pyplot as plt  
  
matplotlib.use('TkAgg')  
  
wines\_data = pd.read\_csv('wine\_Austral.dat', delimiter='\t')  
  
# Преобразование строковых дат в объекты datetime  
months\_days = {'1': 31, '2': 28, '3': 31, '4': 30,  
 '5': 31, '6': 30, '7': 31, '8': 31, '9': 30,  
 '10': 31, '11': 30, '12': 31}  
  
wines\_data['month'] = wines\_data['month\_'].astype(str) + ' ' + wines\_data['year\_'].astype(str)  
months\_list = []  
for m\_date in wines\_data['month']:  
 month\_values = m\_date.split(' ')  
 if isleap(int(month\_values[1])) and month\_values[0] == '2':  
 last\_day = 29  
 else:  
 last\_day = months\_days[month\_values[0]]  
 month\_date = datetime.strptime(str(last\_day) + ' ' + m\_date, '%d %m %Y')  
 months\_list.append(month\_date)  
  
del wines\_data['month']  
del wines\_data['month\_']  
del wines\_data['year\_']  
  
wines\_data.insert(0, 'month', months\_list)  
  
# Выборка столбцов из набора данных  
new\_wine\_data = pd.concat([wines\_data['month'], wines\_data['fort']], axis=1)  
new\_wine\_data.set\_index('month', inplace=True)  
new\_wine\_data.index = pd.to\_datetime(new\_wine\_data.index)  
  
new\_wine\_data = new\_wine\_data.asfreq('m')  
  
decompose = seasonal\_decompose(new\_wine\_data)  
  
f1 = plt.figure()  
f2 = plt.figure()  
f3 = plt.figure()  
# График тренда  
ax1 = f1.add\_subplot(111)  
ax1.plot(decompose.trend, color='blue')  
ax1.set\_title('Тренд')  
ax1.set\_xlabel('Год')  
ax1.set\_ylabel('Литры (в тыс.)')  
ax1.grid(True)  
# График сезонности  
ax2 = f2.add\_subplot(111)  
ax2.plot(decompose.seasonal, color='blue')  
ax2.set\_title('Сезонность')  
ax2.set\_xlabel('Год')  
ax2.set\_ylabel('Литры (в тыс.)')  
ax2.grid(True)  
  
# Прогнозирование - обучение модели sarimax  
learn\_dataset = new\_wine\_data[:]  
  
warnings.simplefilter(action='ignore', category=Warning)  
model = SARIMAX(learn\_dataset, order=(3, 0, 0), seasonal\_order=(0, 1, 0, 12))  
  
result = model.fit()  
  
# Предсказываем поведение ряда на последующие 8 месяцев  
start = len(learn\_dataset)  
end = len(learn\_dataset) + 8  
  
predictions = result.predict(start, end)  
print('\n' + 'Спрогнозированные на 8 месяцев значения:')  
print(predictions)  
  
# Исходный ряд и значения прогноза  
ax3 = f3.add\_subplot(111)  
ax3.plot(new\_wine\_data, color='blue')  
ax3.plot(predictions, color='red')  
ax3.set\_title('Потребление крепленого вина')  
ax3.set\_xlabel('Год')  
ax3.set\_ylabel('Литры (в тыс.)')  
ax3.grid(True)  
plt.show()

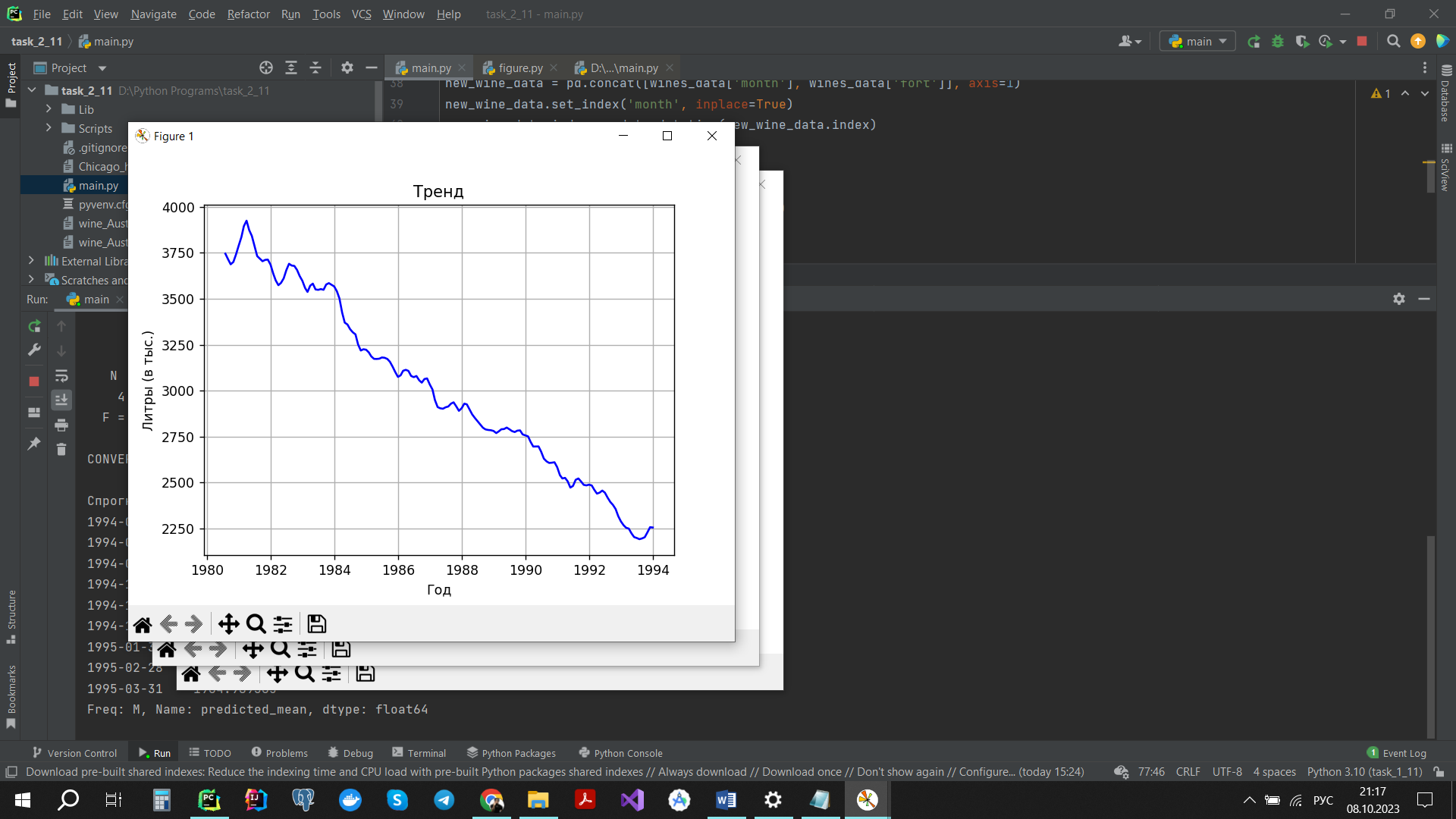
**Результаты работы программы:**

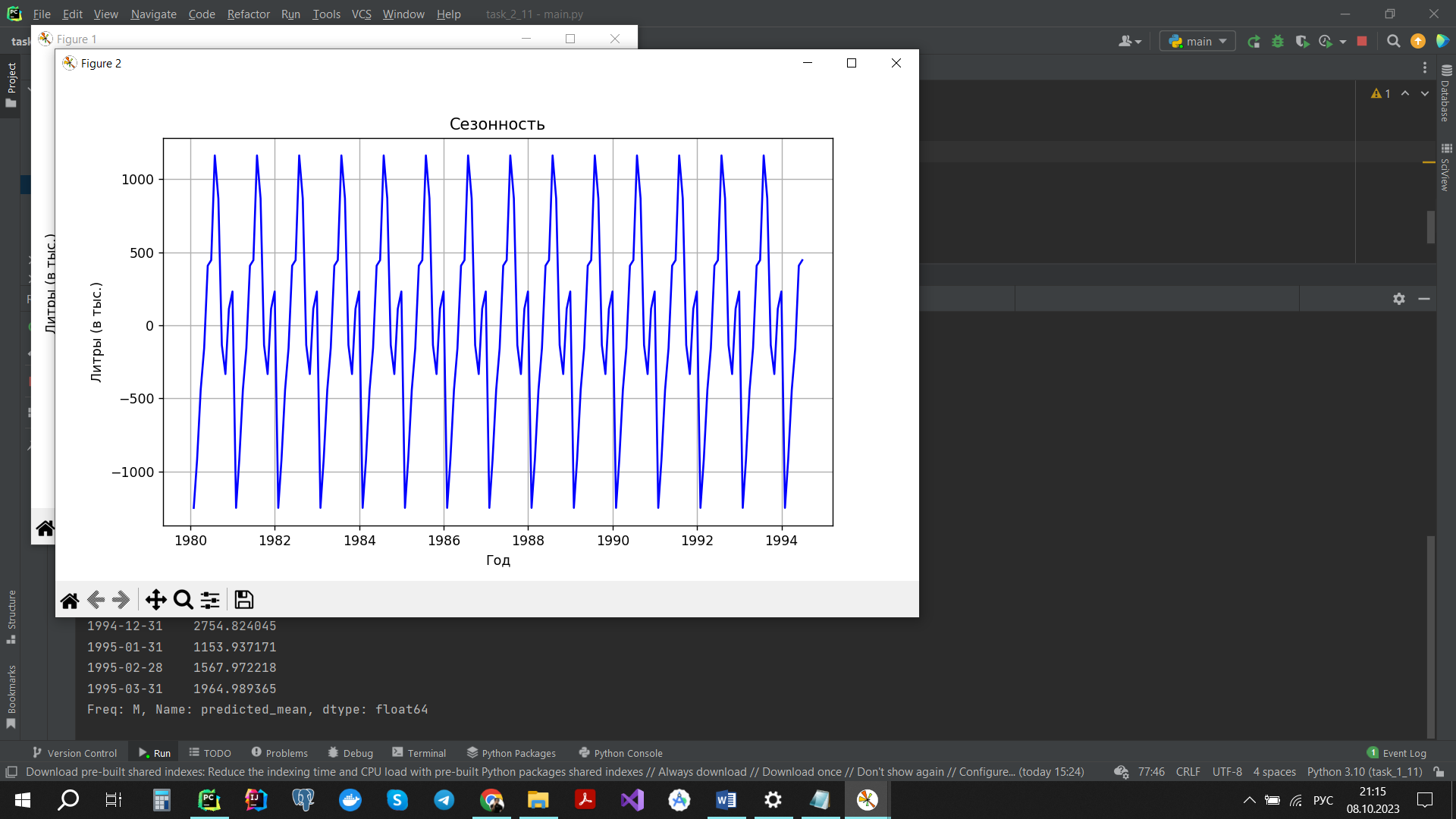
После отработки кода имеем несколько графиков.

Сначала выводится исходный временной ряд, а также прогноз на следующие 8 месяцев.



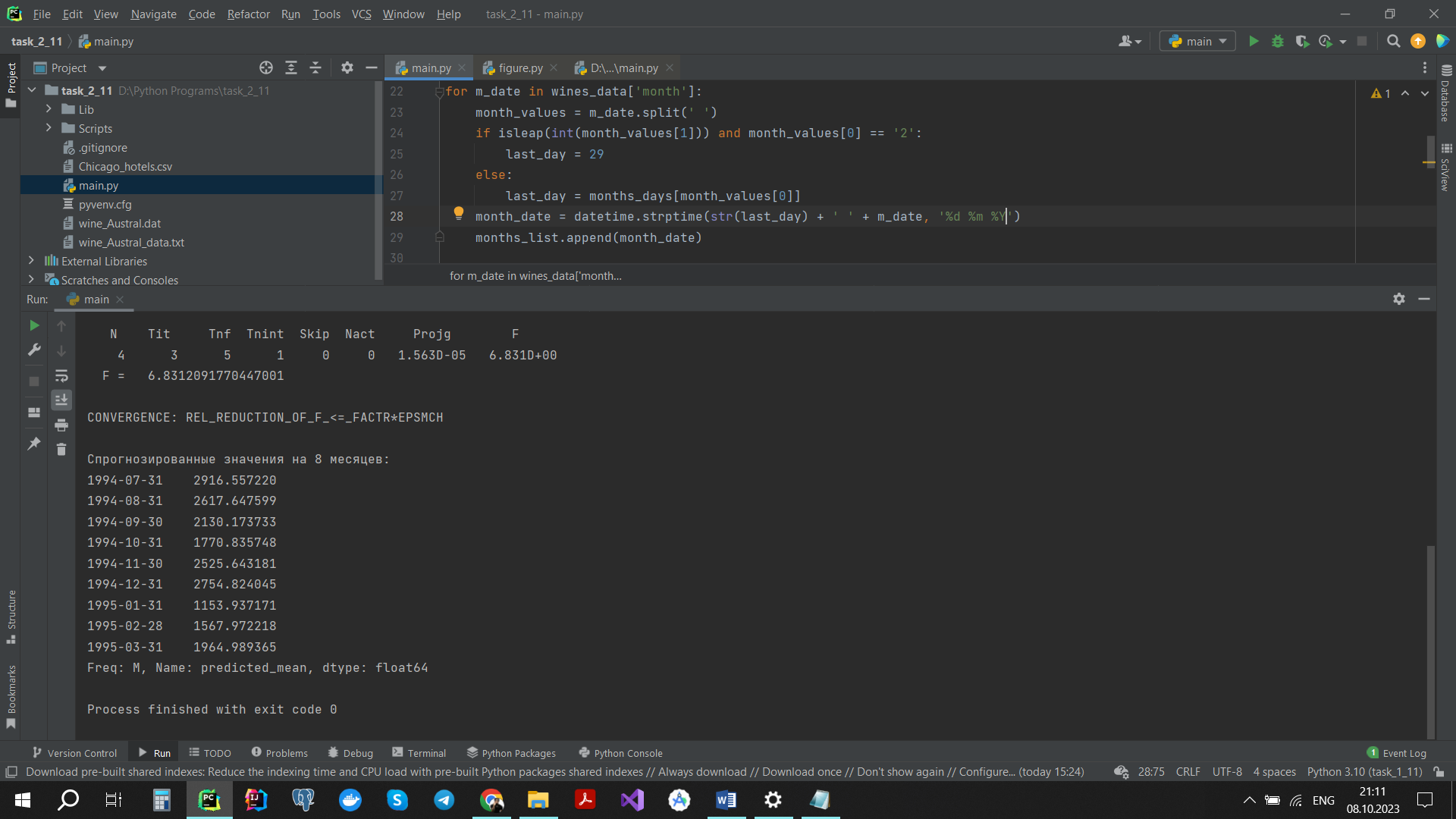
После этого выведутся графики для сезонности и тренда временного ряда.





Как можно заметить у графика нисходящий тренд и четкая сезонность, выбросы незначительны, характер ряда не меняется со временем.

Также в консоль выводятся все спрогнозированные значения рассматриваемой величины для каждого из будущих моментов времени.



**Выводы:**

На основе полученных данных и прогнозов можно проследить тенденцию снижения популярности крепленых вин в Австралии, скорее всего граждане этой страны со временем стали отдавать предпочтение другим видам вин, либо совсем отказались от употребления вина.

**Использованные функции/библиотеки:**

Pandas – хранение csv файла в удобном формате

Matplotlib – визуализация графиков, дендрограмм и многомерного шкалирования

Statsmodels – предоставляет возможность оценить множество различных моделей, а также провести статистические тесты и исследования данных