

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

Воронеж 2023

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)

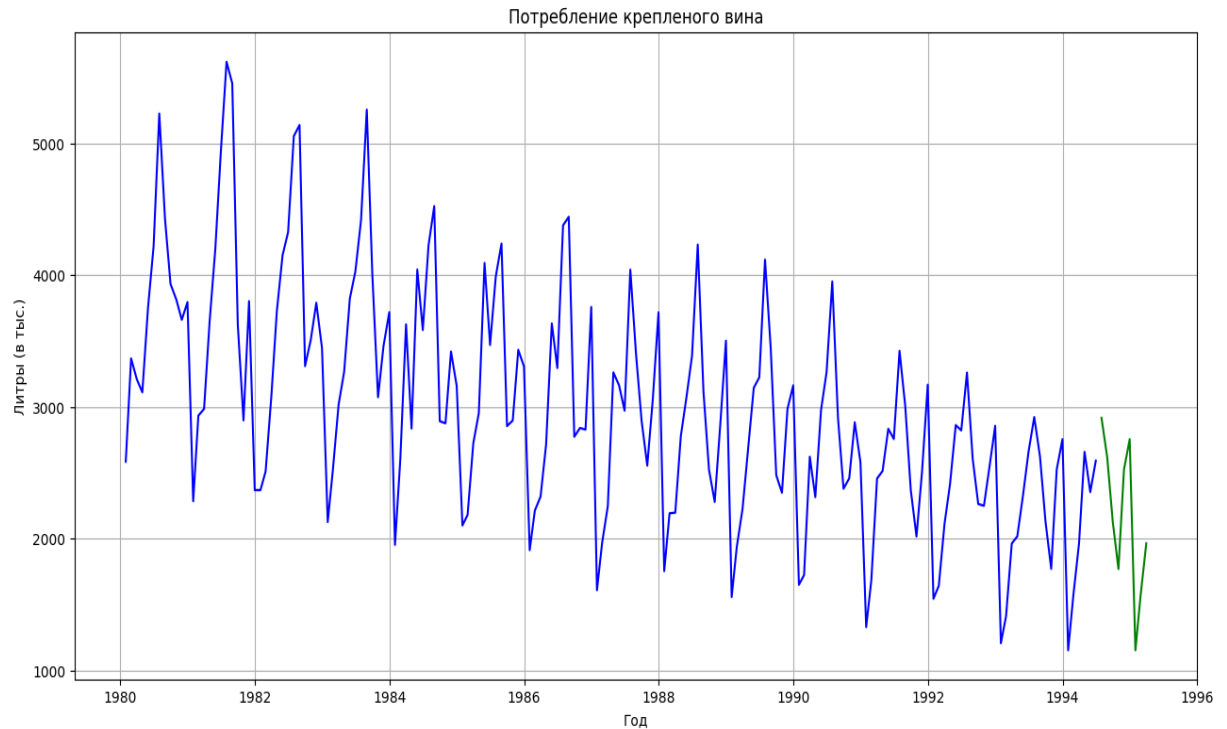
# Исходный ряд и значения прогноза
ax3 = f3.add_subplot(111)
ax3.plot(new_wine_data, color='blue')
ax3.plot(predictions, color='red')
ax3.set_title('Потребление крепленого вина')
ax3.set_xlabel('Год')
ax3.set_ylabel('Литры (в тыс.)')
ax3.grid(True)
plt.show()

```

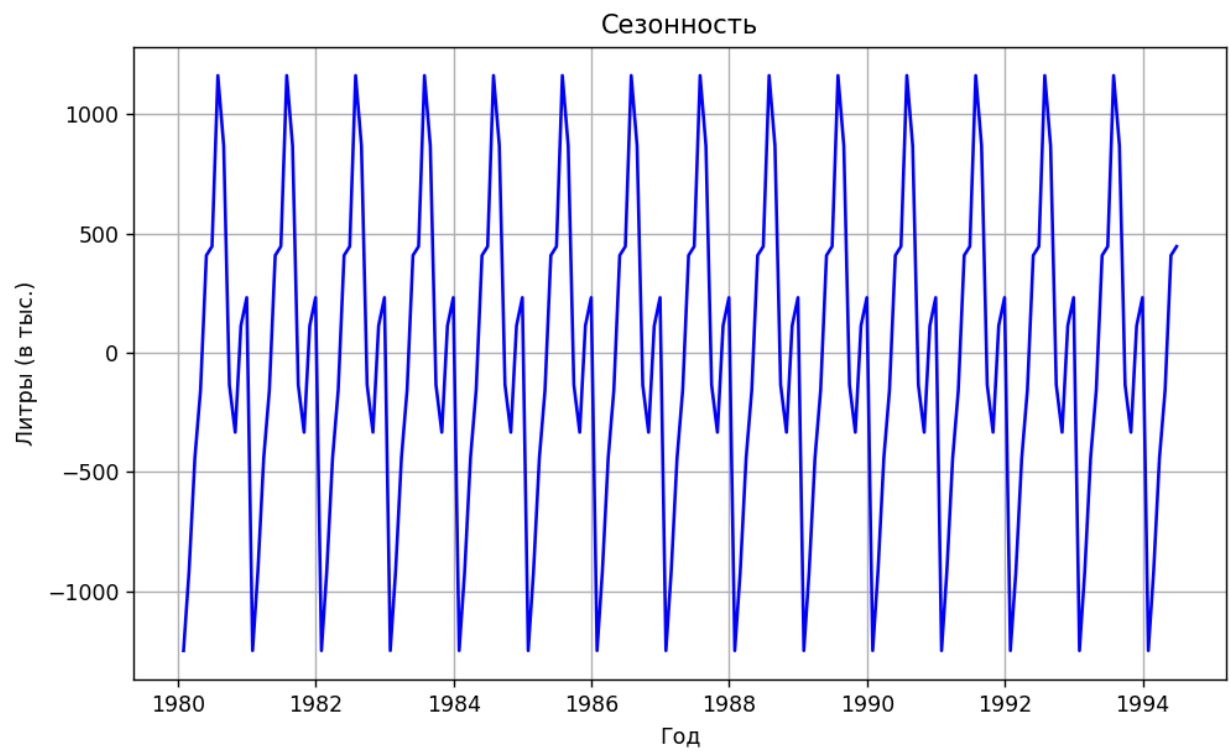
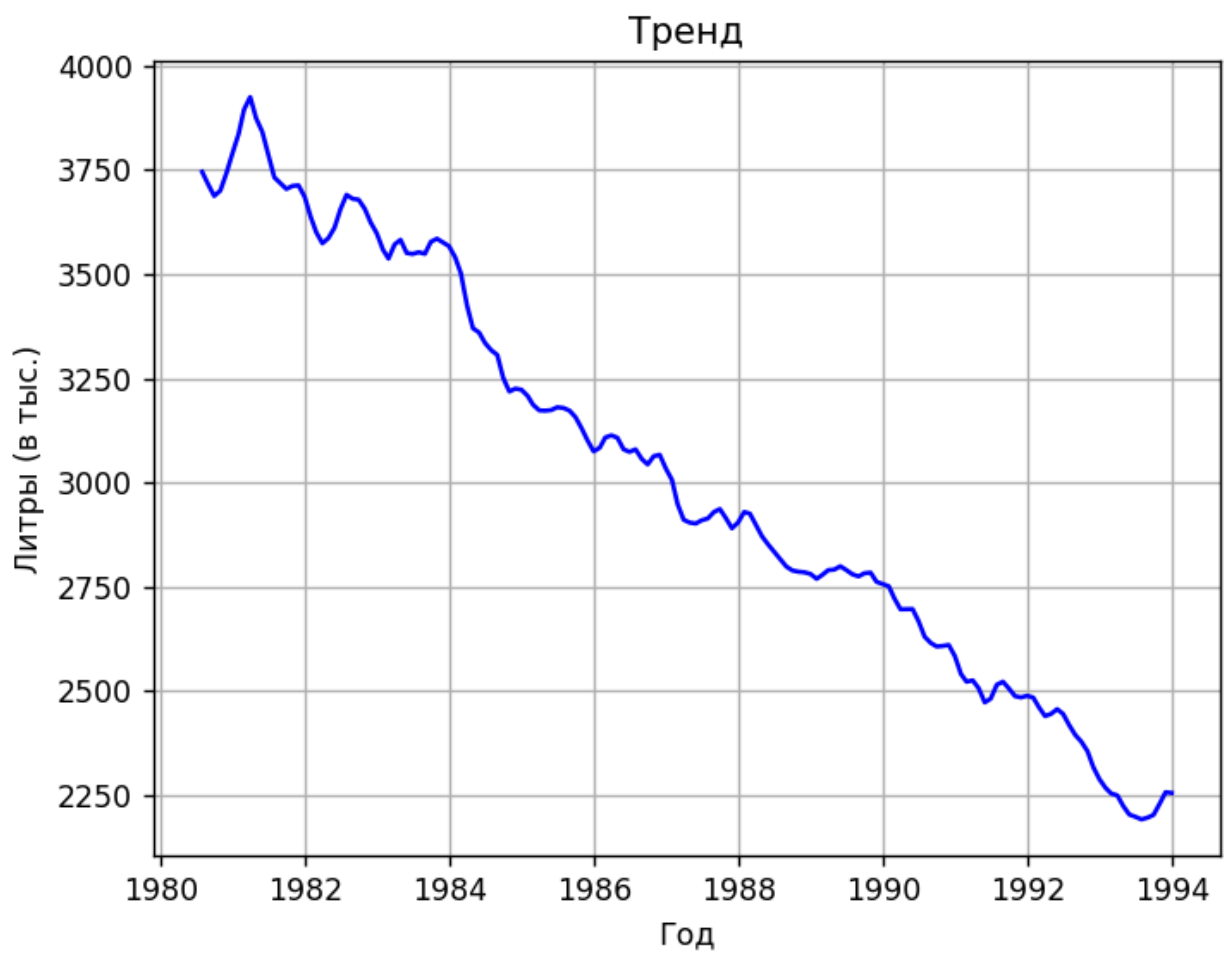
Результаты работы программы:

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.

Спрогнозированные значения на 8 месяцев:	
1994-07-31	2916.557220
1994-08-31	2617.647599
1994-09-30	2130.173733
1994-10-31	1770.835748
1994-11-30	2525.643181
1994-12-31	2754.824045
1995-01-31	1153.937171
1995-02-28	1567.972218
1995-03-31	1964.989365

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