Лабораторная работа №8

**«Анализ временных рядов и прогнозирование».**

Файл «airline-passengers.csv» содержит датасет, представляющий собой месячное отслеживание количества пассажиров в международных авиалиниях, выраженное в тысячах, за период с 1949 по 1960 годы. Ваша задача состоит в том, чтобы провести всесторонний анализ этих данных и сделать прогнозы на будущее.

1. Проведите предварительный анализ данных:

* Загрузите данные и проверьте их на наличие пропусков и аномалий.
* Постройте график временного ряда и визуально оцените динамику числа пассажиров.

2. Разложение временного ряда:

* Разложите временной ряд на тренд, сезонную компоненту и остаток.
* Оцените влияние каждого компонента на общий уровень количества пассажиров.

В качестве тренда выберите один из вариантов, в соответствии с номером по списку (1 – линейный, 2 – параболический, 3 – из экспоненциального, показательного или логарифмического тот, у которого лучше точность аппроксимации).

Для определения сезонной компоненты чётные варианты применяют ряды Фурье, нечётные – индексы сезонности (константные и динамические, оценив качество обоих подходов).

3. Прогнозирование:

* Сделайте точечный прогноз числа пассажиров на следующий 1961 год.
* Проверьте качество подхода на основе метрик ошибки (например, RMSE, MAE), сделав прогноз на 1960 год на основе предыдущих значений.
* Представьте результаты в виде графика, где отображены реальные и прогнозируемые значения.
* Желающие могут сделать интервальный прогноз, для 95% и 98% уровня значимости, отразив результаты в табличном и графическом виде (+2 балла).

4. Проверка стационарности:

* Проверьте временнóй ряд на стационарность (любым из методов).
* Если ряд нестационарен, примените методы преобразования к стационарному виду (например, дифференцирование).

5. Использование подхода ARIMA (SARIMA):

* Постройте и исследуйте автокорреляционную функцию (ACF), чтобы определить подходящие параметры для модели ARIMA (SARIMA).
* Сделайте прогноз на 1961 год и тестирование модели на 1960 год.
* Обсудите полученные результаты и предложите возможные причины отклонений между реальными данными и прогнозами.
* Предложите рекомендации по улучшению точности прогнозов.

6. Подведите итоги проведенного анализа и сформулируйте выводы.

Задание должно быть выполнено в формате Jupyter Notebook.

Код должен быть хорошо структурирован и прокомментирован.

Графики должны быть четкими и информативными, с соответствующими подписями и легендами.

Выводы и обсуждение должны быть представлены в письменной форме и содержать обоснованные аргументы.

Срок выполнения задания – одна неделя (6 баллов).

Программа защищается на паре.

Можно работать в группе по 2 человека.