# Аналитическая записка\_15

## Тема

Построение и оценка эконометрической модели зависимости индекса реального ВВП от индекса реальной зарплаты.

## Цель анализа

Исследовать взаимосвязь между индексом реальной зарплаты (Real\_Wage\_Index) и индексом реального ВВП (Real\_GDP\_Index), оценить спецификацию модели, проверить выполненные предпосылки и построить прогнозное значение с доверительным интервалом.

## Задание 1: Построение спецификации эконометрической модели

Результаты:  
- Коэффициент корреляции между индексом реальной зарплаты и индексом реального ВВП составил 0.7483, что указывает на умеренную положительную связь между переменными.  
- p-значение: 1.55e-07, что значительно ниже уровня значимости 0.05, подтверждая статистическую значимость корреляции. Это указывает на высокую вероятность того, что связь между переменными не случайна.

## Задание 2: Оценивание качества спецификации модели

Построение модели: Была построена линейная регрессионная модель, описывающая зависимость индекса реального ВВП от индекса реальной зарплаты. Результаты оценки параметров модели:  
  
- Константа: 49.78, p-значение = 0.005, что указывает на статистическую значимость константы на уровне 0.05.  
- Коэффициент при Real\_Wage\_Index: 0.4665, p-значение < 0.0001, что подтверждает, что индекс реальной зарплаты оказывает статистически значимое влияние на индекс реального ВВП.  
- Коэффициент детерминации (R^2): 0.560, что означает, что модель объясняет 56.0% изменчивости индекса реального ВВП на основе индекса реальной зарплаты.  
  
Средняя относительная ошибка аппроксимации составила 5.04%, что подтверждает достаточную точность модели.  
  
Вывод: Модель продемонстрировала умеренную объяснительную способность, так как значительная часть изменчивости ВВП объясняется изменениями в индексе реальной зарплаты.

## Задание 3: Проверка остатков на нормальное распределение

Методология: Для проверки нормальности распределения остатков использован тест Шапиро-Уилка. Этот тест выбран, так как он хорошо подходит для небольших и средних выборок и чувствителен к отклонениям от нормальности.

Для проверки нормальности распределения остатков выбран **тест Шапиро-Уилка**. Основные причины его использования заключаются в следующем:

1. **Подходит для малых и средних выборок**: Тест Шапиро-Уилка считается одним из наиболее мощных тестов на нормальность для небольших и средних выборок, особенно когда количество данных не очень велико. Он оптимизирован для этих случаев, что делает его более надежным в сравнении с другими тестами.
2. **Высокая чувствительность к отклонениям от нормальности**: Шапиро-Уилк имеет высокую чувствительность к различным типам отклонений от нормального распределения. Это важно при анализе остатков, так как даже небольшие отклонения от нормальности могут повлиять на выводы о корректности модели.
3. **Частое применение в эконометрическом анализе**: Этот тест является стандартом для проверки нормальности остатков в эконометрии и статистике. Его результаты легко интерпретировать, и он широко поддерживается в аналитических пакетах, таких как Python и R.

Результаты:  
- Статистика теста Шапиро-Уилка: 0.9440  
- p-значение: 0.0679  
  
Так как p-значение больше 0.05, у нас нет оснований отвергать гипотезу о нормальности распределения остатков. Это означает, что остатки модели можно считать нормально распределёнными.  
  
Вывод: Остатки нормально распределены, что подтверждает выполнение предпосылки нормальности для модели.

## Задание 4: Проверка предпосылки теоремы Гаусса-Маркова о гомоскедастичности

Методология: Для проверки гомоскедастичности проведён тест Бреуша-Пагана, который проверяет постоянство дисперсии остатков. Этот тест выбран, так как он эффективно выявляет гетероскедастичность в линейных моделях.

Для проверки гомоскедастичности выбран **тест Бреуша-Пагана**, потому что:

1. **Прямо тестирует гетероскедастичность** — проверяет, меняется ли дисперсия остатков в зависимости от независимых переменных.
2. **Подходит для линейных моделей** — часто используется для оценки соответствия остатков предпосылке постоянной дисперсии.
3. **Удобен для интерпретации** — результат дает статистику и pp-значение, что позволяет легко определить наличие или отсутствие гетероскедастичности.

Тест Голдфелда-Квандта не использован, так как он менее универсален и предполагает влияние только одной переменной на гетероскедастичность.

Результаты:  
- Статистика Бреуша-Пагана: 0.5633  
- p-значение: 0.4530  
  
Так как p-значение больше 0.05, у нас нет оснований отвергать гипотезу о гомоскедастичности. Это означает, что дисперсия остатков постоянна, и гетероскедастичности не наблюдается.  
  
Вывод: Гомоскедастичность остатков подтверждается, что соответствует предпосылкам линейной регрессии.

## Задание 5: Построение прогнозного доверительного интервала

Цель: Рассчитать прогнозное значение индекса реального ВВП при уровне индекса реальной зарплаты, равном 105% от её среднего значения, а также построить доверительный интервал для этого прогноза.  
  
Результаты:  
- Среднее значение Real\_Wage\_Index: примерно 246.9  
- 105% от среднего значения Real\_Wage\_Index: 259.2  
- Точечное прогнозное значение Real\_GDP\_Index при Real\_Wage\_Index = 259.2 составляет 164.16.  
- Доверительный интервал (95%) для прогнозного значения: [160.66, 167.67].  
  
Вывод: Прогнозируемое значение индекса реального ВВП при увеличении индекса реальной зарплаты на 5% от среднего уровня составляет 164.16. Доверительный интервал на уровне значимости 0.05 показывает, что при данном уровне индекса реальной зарплаты ВВП с вероятностью 95% будет находиться в диапазоне от 160.66 до 167.67.

## Задание 6:

1. **Коэффициент корреляции Пирсона**:
   * Коэффициент корреляции между индексом реальной зарплаты (WAG\_Q) и индексом реального ВВП РФ (GDPEA\_Q\_DIRI) составляет **0.7483**, что указывает на умеренную положительную связь между переменными.
   * pp-значение для корреляции равно **1.55e-07**, что значительно меньше уровня значимости 0.05, подтверждая статистическую значимость корреляции. Это указывает на высокую вероятность того, что выявленная связь не случайна.
2. **Модель линейной регрессии**:
   * **Константа**: 49.78, pp-значение = 0.005, что указывает на статистическую значимость константы на уровне 0.05.
   * **Коэффициент при WAG\_Q**: 0.4665, pp-значение < 0.0001, что подтверждает, что индекс реальной зарплаты оказывает статистически значимое влияние на индекс реального ВВП РФ.
   * **Коэффициент детерминации R2R2**: 0.560. Это означает, что модель объясняет 56.0% изменчивости индекса реального ВВП на основе индекса реальной зарплаты, что свидетельствует о средней объяснительной способности модели.
3. **Средняя относительная ошибка аппроксимации** составила **5.04%**, что указывает на достаточную точность модели.

**Заключение:**

Модель линейной регрессии показала, что индекс реальной зарплаты (WAG\_Q) оказывает значительное влияние на индекс реального ВВП РФ (GDPEA\_Q\_DIRI). Значимый коэффициент при WAG\_Q и умеренный коэффициент детерминации (0.560) подтверждают, что изменения в реальной зарплате связаны с изменениями в реальном ВВП РФ. Эта модель может использоваться для оценки того, как изменения в реальной зарплате могут влиять на экономическую активность, измеряемую индексом ВВП.

## Общий вывод

В ходе анализа была построена эконометрическая модель, описывающая зависимость индекса реального ВВП от индекса реальной зарплаты. Модель показала умеренную объяснительную способность, а коэффициент при индексе реальной зарплаты оказался значимым. Проверки на нормальность остатков и гомоскедастичность подтвердили выполнение предпосылок регрессионного анализа.  
  
Полученные прогнозные значения и доверительный интервал могут быть полезны для оценки возможного уровня ВВП при изменении индекса реальной зарплаты, что может послужить основой для принятия решений в экономической политике и планировании.