**Федеральное государственное автономное образовательное учреждение**

**высшего образования   
"Национальный исследовательский университет   
"Высшая школа экономики"**

**Московский институт электроники и математики Национального**

**исследовательского университета "Высшая школа экономики" им. А.Н. Тихонова**

**Департамент прикладной математики**

Направление подготовки 01.03.04 «Прикладная математика»

**О Т Ч Е Т**

**по производственной практике**

**Студент** Мартынов Н.И.  **\_\_\_\_\_\_\_**161**\_\_\_\_\_\_\_\_\_**

(Фамилия И.О.) номер группы

**Руководитель практики:** Романенко Е.М. Главный специалист Карьерный центр, АО «ДОМ.РФ»

(Фамилия И.О., должность и место работы, подпись)

**Ответственный за организацию практики:** Внуков А.А., доцент, НИУ ВШЭ

(Фамилия И.О., должность, подпись)

**Практика пройдена с оценкой \_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Дата \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Москва 2019**

**Цели и задачи практики**

**Цель**:

Разработать приложение с графическим пользовательским интерфейсом для построения временных зависимостей различных характеристик рынка недвижимости в регионах РФ.

**Задачи**:

1. Ознакомление с устройством рынка недвижимости в регионах РФ
2. Визуализация данных
3. Исследование возможностей реализации графического пользовательского интерфейса
4. Программная реализация графического пользовательского интерфейса и алгоритма отрисовки зависимостей характеристик рынка
5. Составление отчета о проделанной работе

**Постановка задачи**

Имеется массив данных о заключённых сделках по недвижимости с различными характеристиками объектов по различным регионам РФ. Необходимо проанализировать данные, вычислить основные характеристики, реализовать на ЯП Python приложение для графической визуализации зависимостей характеристик и написать алгоритм по вычислению массивов данных для этих зависимостей.

**Визуализация данных**

Для более детального и точного анализа можно разбить данные на кластеры, если такое разбиение существует. Для того, чтобы оценить шансы существования таких разбиений применялись алгоритмы PCA и t-SNE, после чего преобразованные данные отображались в трехмерном пространстве или на плоскости с помощью matplotlib. Кроме того, визуализировались основанные статистики данных вдоль одной из размерностей данных и также динамика этих статистик по размерностям (медиана, среднее, дисперсия, центральные моменты и тд)

**Описание решения поставленной задачи**

Для написания приложения и, в частности, GUI использовался Python Qt (PyQt). Пользовательский интерфейс обеспечивает выбор характеристик для выбора среза данных, после чего по выбранному срезу строятся выбранные пользователем статистики.

Для выбора среза по данным и для удобства хранения данных использовалась библиотека pandas и входящие в неё функции, а готовые зависимости отображались с помощью matplotlib.

Для быстрого вычисления характеристик данных использовалась библиотека numpy.

**Полученные результаты**

Было получено приложение с GUI, которое представляет собой программу на ЯП Python, приложение можно запускать из терминала, после чего выбирать срез данных, по которому нужно построить статистики, которые пользователь также выбирает из предложенного списка.

Продукт имеет коммерческое значение.

**Выводы**

Средства ЯП Python обеспечивают быстрое и эффективное решение задачи по написанию GUI, а также позволяют необходимым образом обрабатывать массивы данных и вычислять необходимые величины.