

## **Аннотация рабочей программы дисциплины «Основы анализа данных на Python»**

**Целью** освоения дисциплины «Основы анализа данных на Python» является формирование комплекса знаний о средствах программирования на Python и специализированных библиотек анализа данных, а так же применение полученных знаний для целей бизнес-аналитики.

### **Задачи дисциплины:**

- формирование теоретических знаний и практических навыков по основам анализа данных в Python
- формирование навыков использования Python для работы с табличными данными, импортируемыми или экспортируемыми популярными программами для анализа данных
- приобретение практических навыков использования набора библиотек языка Python для решения задач в сфере бизнес-аналитики

### **Тема 1. Введение в анализ данных на языке Python**

1. Python – стандарт для работы с большими данными
2. Прикладные задачи бизнес-аналитиков, для решения которых применяется язык Python

### **Тема 2. Основные операторы и типы данных**

1. Работа со строками
2. Циклы for
3. Преобразования между типами
4. Работа с последовательностями
5. Работа с текстовыми файлами
6. Работа с модулями

### **Тема 3. Обмен данными между Python и популярными табличными приложениями**

1. Формат табличных данных CSV
2. Выгрузка данных для обработки в Python
3. Открытие CSV в Python
4. Формирование файла CSV средствами Python
5. Загрузка сформированных в Python данных в программу
6. Формат табличных данных с фиксированной шириной ячеек (FWF)
7. Прямая интеграция с документами Excel

### **Тема 4. Использование регулярных выражений (regex)**

1. Описание набора символов
2. Описание количества символов
3. Начало и конец строки
4. Степень детализации регулярных выражений
5. Группы в регулярных выражениях
6. Библиотека для регулярных выражений в Python

### **Тема 5. Выгрузка данных из онлайн API**

1. Отправка запросов из Python
2. Формат иерархических данных JSON
3. Формат иерархических данных XML

### **Тема 6. Обзор модулей Python для анализа данных**

1. Модуль для работы с числовыми данными NumPy
2. Модуль для работы с табличным представлением данных Pandas
3. Обзор библиотеки Matplotlib для визуализации данных в Python

**Объем дисциплины 3 з.е.**

**Форма промежуточного контроля – зачет.**