ФИО студента \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Дисциплина: Анализ данных

Задание 1

Откройте файл **Worldbank\_data.dta**

Проведите кластерный анализ данных с использованием всех переменных кроме year, groups, country\_name, country\_code. Используйте метод Варда.

1. Определите число кластеров, объясните своё решение.

2. Создайте дендрограмму, скопируйте её в этот файл и опишите её.

3. Опишите получившиеся кластеры с помощью описательной статистки. Нужно не просто скопировать таблицы, но детально описать полученные результаты.

4. Дайте названия кластерам.

5. Сохраните переменную принадлежности к кластеру в файле.

Задание 2

Откройте файл **video\_games\_sales.dta**

Проведите кластерный анализ данных с использованием не менее пяти интервальных переменной. Используйте метод k-средних.

1. Определите число кластеров, объясните своё решение.

2. Опишите получившиеся кластеры с помощью описательной статистки. Нужно не просто скопировать таблицы, но детально описать полученные результаты.

3. Дайте названия кластерам.

4. Сохраните переменную принадлежности к кластеру в файле.

Загрузите, пожалуйста, этот файл с ответами и изменённые файлы с данными в MS Teams в раздел «Задания».