**Задание для Аналитика**

На основании вопросов и таблицы с данными, представленными в Excel-файле «Задание для Аналитика», необходимо подготовить небольшой аналитический отчет (не менее 3-х страниц) в формате Word. В Excel-файле «Задание для Аналитика» приведены данные опроса граждан (1500 респондентов) Республики Таджикистан за 2019 год.

Для выполнения задания необходимо:

1. Определить **количество и долю опрошенных граждан** по регионам, типу местности (городская местность/сельская местность) и полу респондентов. Построить диаграмму и описать результаты.
2. Построить таблицу по **основным источникам информации в нашей стране** целом, а также в разрезе регионов и типа местности (городская/сельская). Данные показать в процентах.
3. Определить доли респондентов, которые **выходят в интернет через компьютер, мобильный телефон и VPN** в разрезе регионов, типа местности и пола респондентов. Подготовить таблицы и диаграммы, и описать результаты.
4. Определите, какими **социальными сетями, чатами и мессенджерами** в основном пользуются респонденты в целом по стране, мужчины и женщины, по регионам.
5. Определите долю респондентов, у которых/членов семьи которых есть:

* сотовый телефон;
* компьютер;
* спутниковое телевидение;
* автомобиль.

Постройте линейчатые диаграммы, указав доли в целом по стране и по регионам. Опишите результаты.

1. Выявить **средний возраст респондентов** (всех, а также мужчин и женщин по-отдельности). Разбить возраст респондентов на следующие возрастные категории:

* 18 – 24 года;
* 25-30 лет;
* 31-35 лет;
* 36-40 лет;
* 41-45 лет;
* 46-50 лет;
* 51-55 лет;
* 56-60 лет;
* 61-65 лет;
* Старше 65 лет.

Выявить **долю каждой возрастной категории** (по всей республике, по полу и регионам).

1. В базе данных есть еще другие показатели (**образование, семейное положение респондентов**). Проявите творческий подход, проанализируйте и опишите эти показатели.

Задание можно выполнить не по порядку. При описании результатов используйте таблицы и диаграммы, которые построили в Excel. На выполнение этого задания выделено 3 дня. Сохраните все свои рабочие документы. Приветствуется использование нестандартных и творческих подходов/методов к анализу, иллюстрации и подаче результатов.

**Результат работы**

В ходе работы был проведен исследовательские анализ с помощью набор данных граждан (1500 респондентов) Республики Таджикистан за 2019 год.

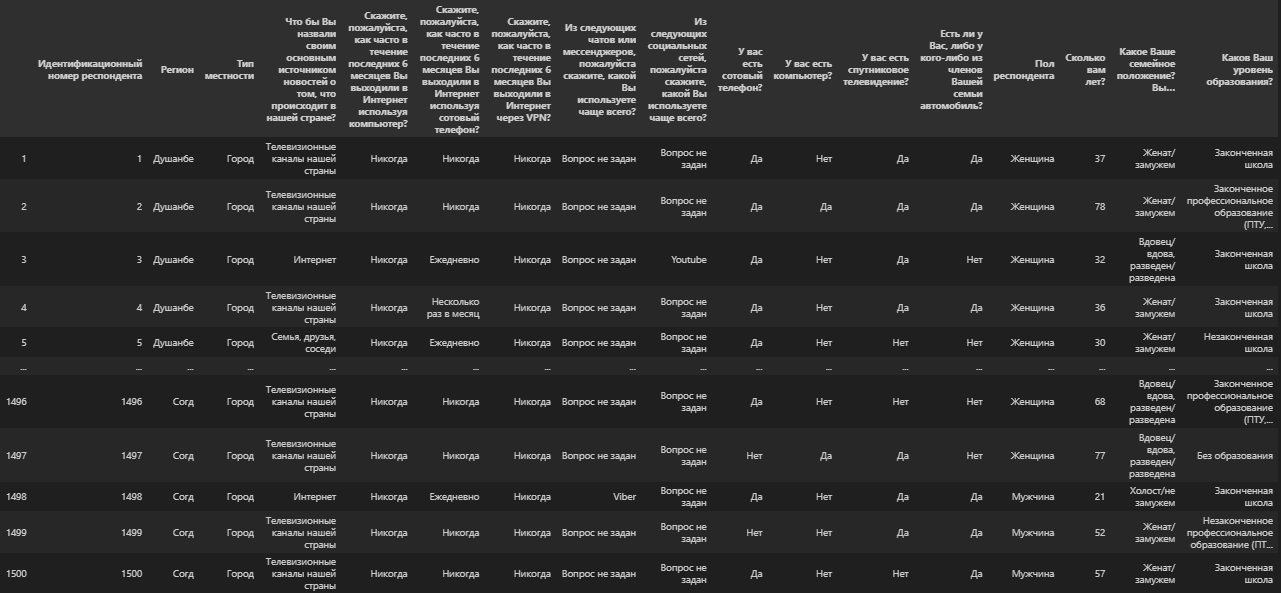
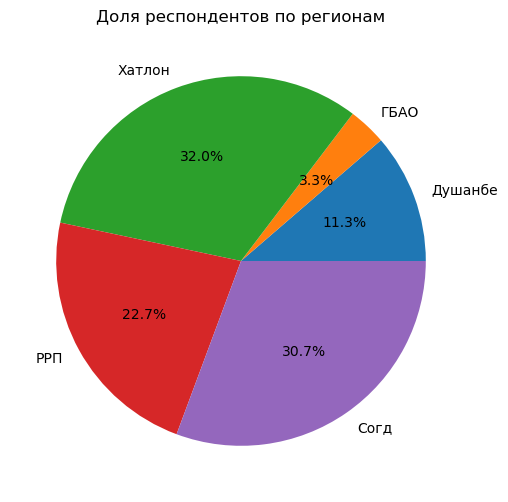


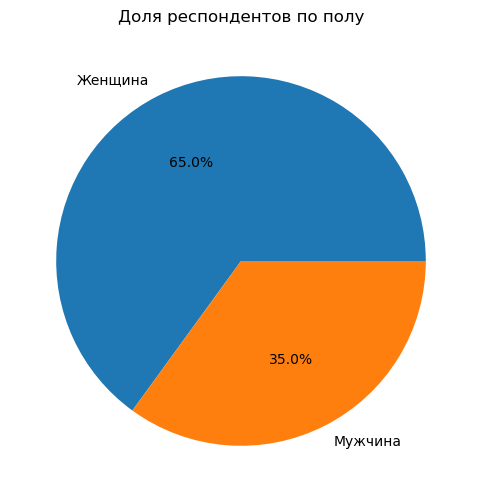
Рисунок 1. Набор данных граждан (1500 респондентов) Республики Таджикистан за 2019 год

Был использовано язык программирования Python с помощью симулятор Jupyter Notebook.

**Ответ к 1 заданий**

Производится анализ данных по регионам, типу местности и полу респондентов. Результаты представляются в виде доли респондентов в процентах для каждой категории, а также визуализируются с помощью круговых диаграмм.

Круговая диаграмма показывает процентное соотношение респондентов в разных регионах. Высокие доли респондентов в определенных регионах могут указывать на особые тенденции или предпочтения, связанные с этими регионами.

****

Доля респондентов по полу: Круговая диаграмма показывает процентное соотношение респондентов по полу.

Доля респондентов по типу местности: Круговая диаграмма отображает процентное соотношение респондентов в разных типах местности. Это может помочь в понимании различий во взглядах и предпочтениях между городскими и сельскими резидентами

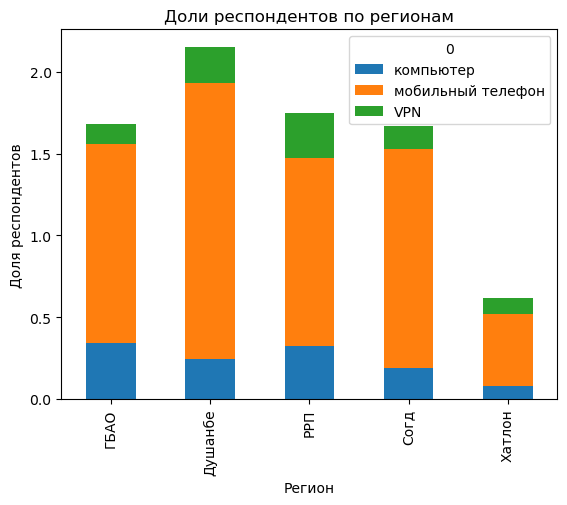
**Ответ к 2 заданий**

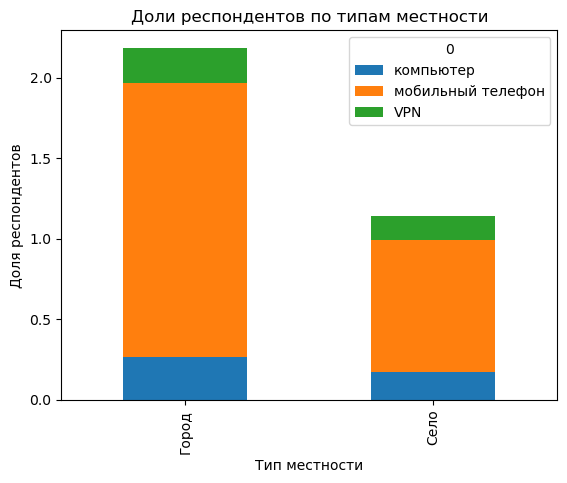


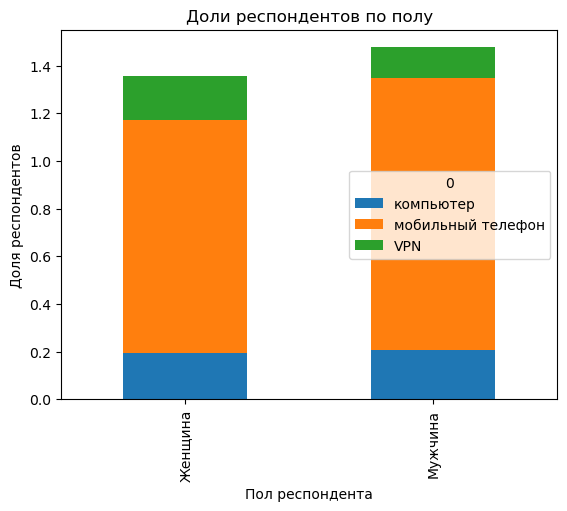
Рисунок 2. Сводная таблица, отражающая информацию по 'Регион', 'Тип местности', 'Источник информации' и 'percentage

Соответствии поставленные задач нам было необходимо построить создать сводной таблицу которое бы отражал нам процентной показатель источник информации по агрегации регион и тип местности. Вышеуказанной рисунок является ожидаемым результатом.

**Ответ к 3 заданий**

Чтобы определить доли респондентов, которые выходят в интернет через компьютер, мобильный телефон и VPN. самого начало группируем данных по регионам и вычисляем долю и так будем группировать, и вычислять для типа местности и полу респондентов. После получение результаты можно построить диаграмму. Результат работы в нижних рисунках указано.





**Ответ к 4 заданий**

У нас есть столбец "социальные сети”, “чаты и мессенджеры" в базе данных, содержащей информацию о предпочтениях респондентов, требуется определить количество использования со стороны респондентов в разрезе регионов, пола и типа местности.

Построим сводной таблицу по группировки Регион и Пол респондента затем вычислим количество респондентов по социальным сетям и по чаты и мессенджеры.

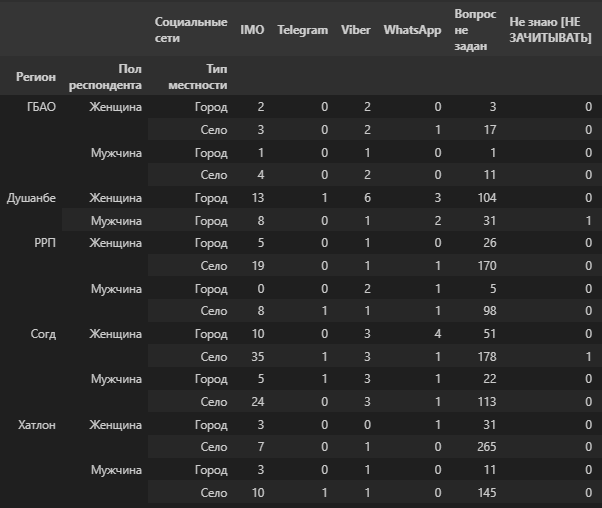


Рисунок 3. Группировка данных по социальным сетям в разрезе регионов, пола и типа местности

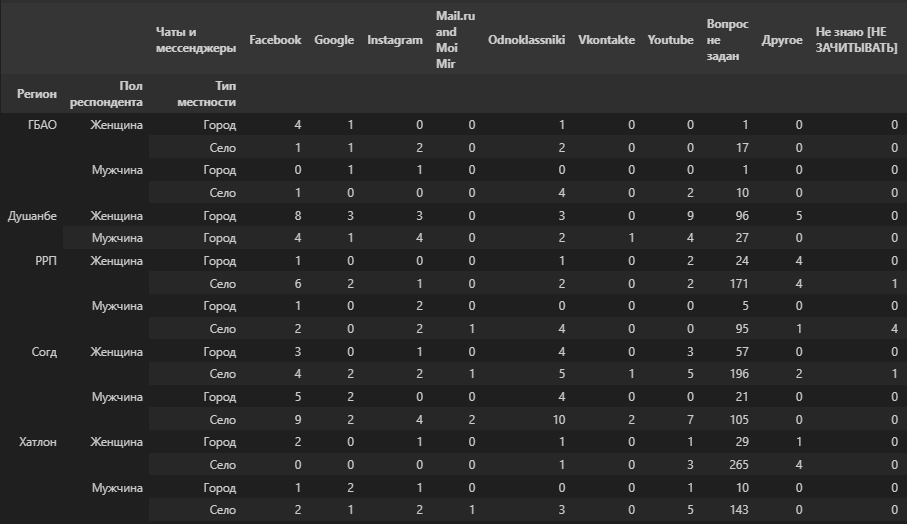


Рисунок 4. Группировка данных по чатам и мессенджерам в разрезе регионов, пола и типа местности

**Ответ к 5 заданий**

Определяем доли респондентов с указанными предметами (Сотовый телефон, Компьютер, Спутниковое телевидение, Автомобиль) затем группируем данных по регионам построим линейчатых диаграмму.

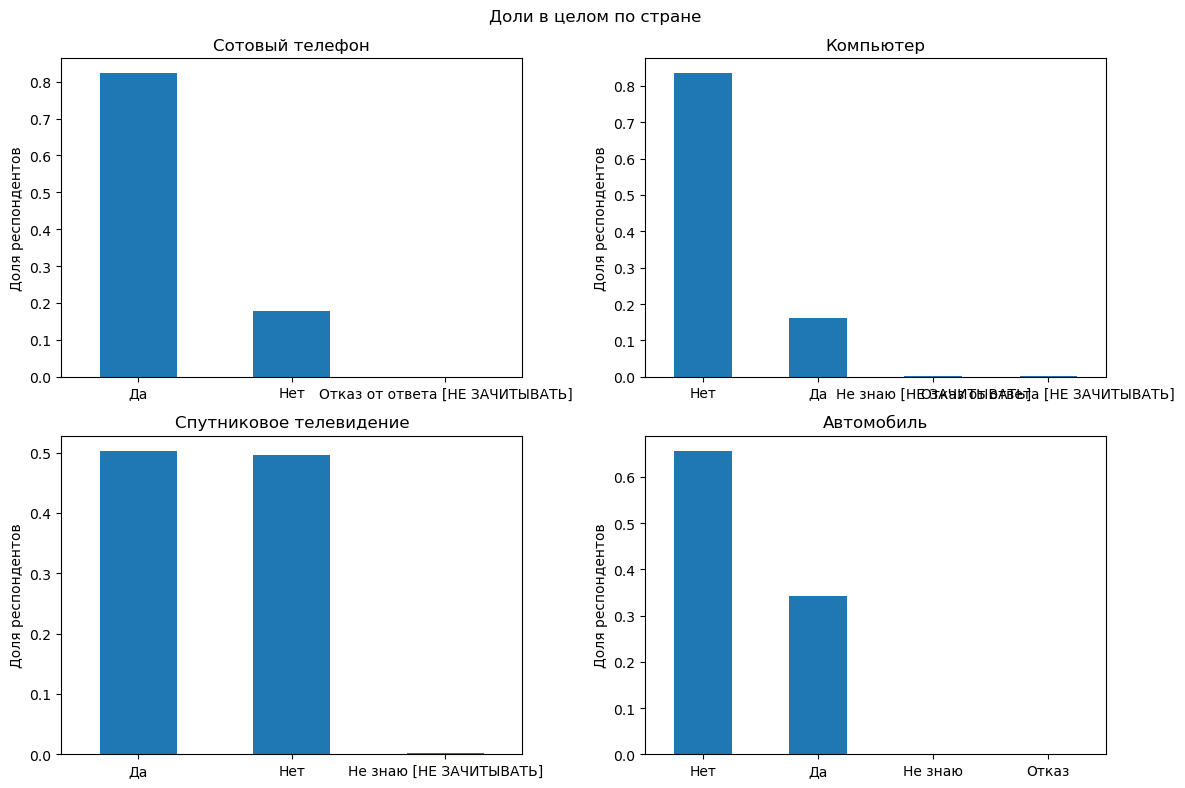


Рисунок 5. Доли в целом по стране

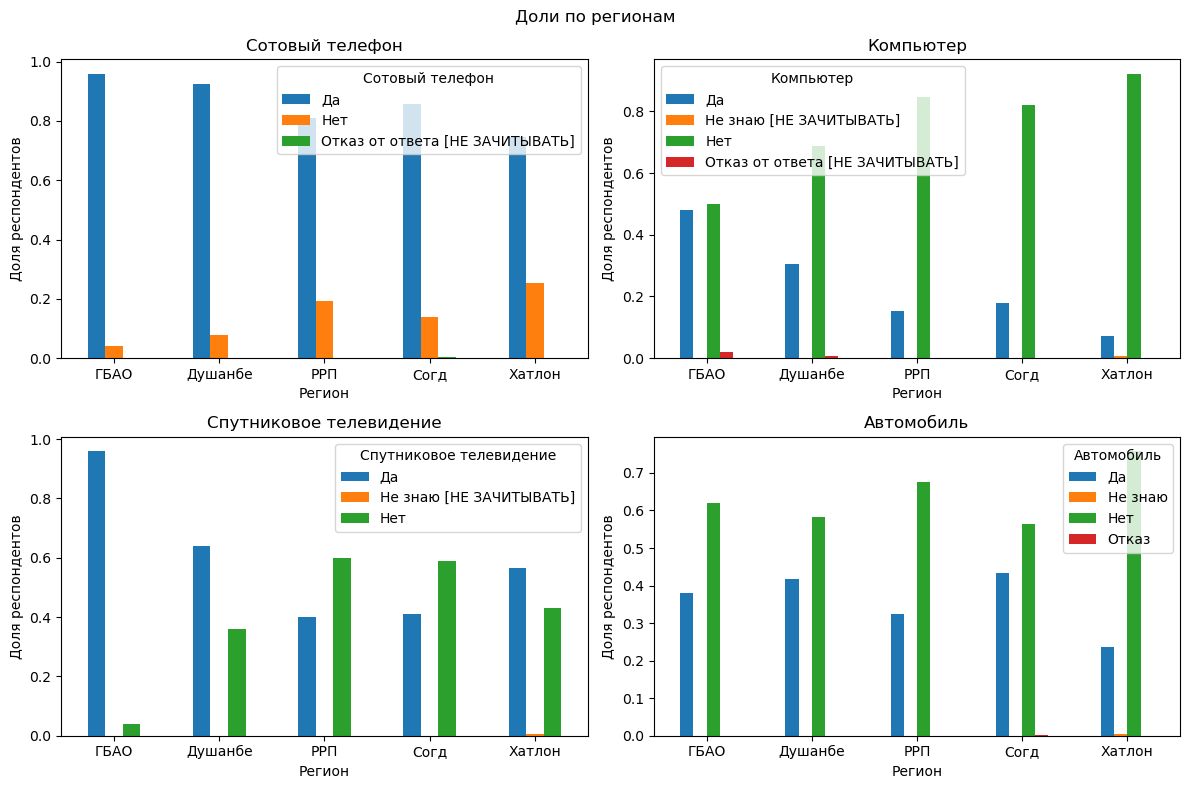
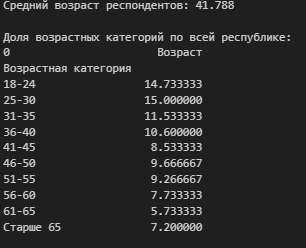
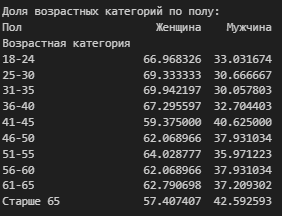
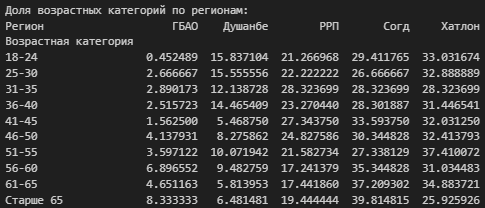
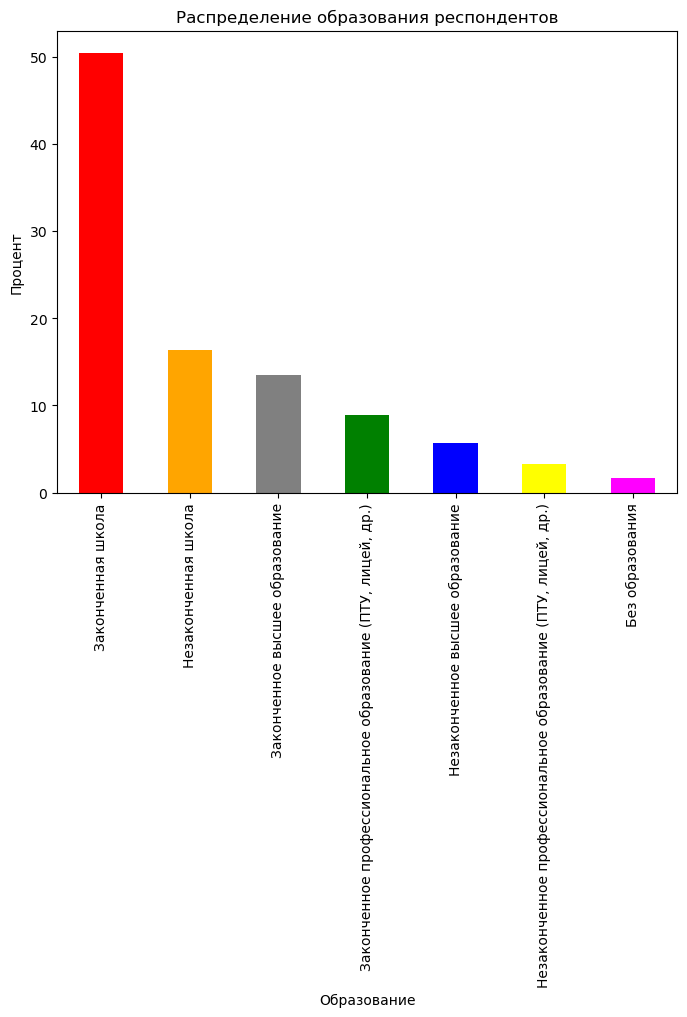


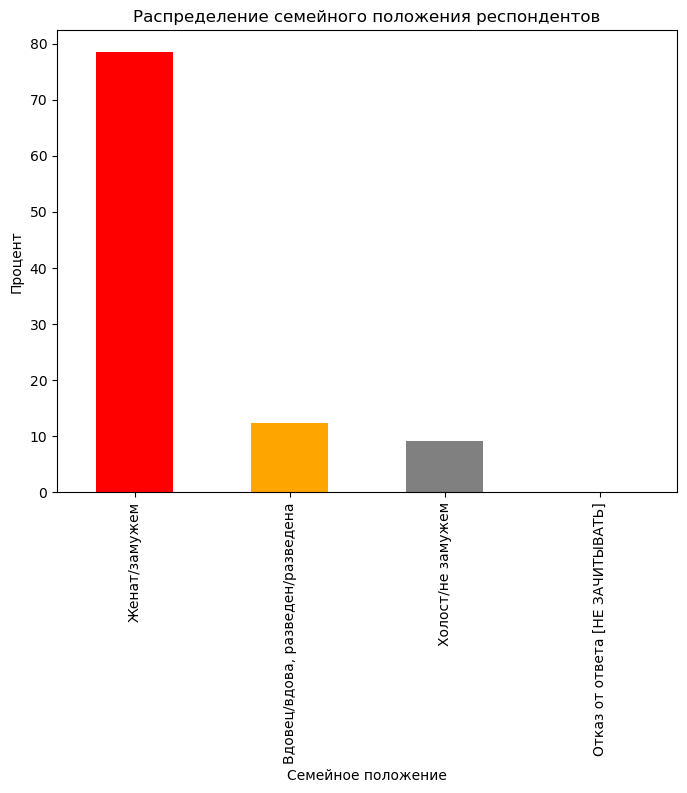
Рисунок 6. Доли по регионам

**Ответ к 6 заданий**



**Ответ к 7 заданий**

****

****