

## Практическое задание

### Часть I

- Задание включает в себя описание ситуации, файл данных.
- Описание ситуации и вопросы одинаковы для всех слушателей.
- Каждый вопрос имеет максимальную оценку в баллах, которая указана в скобках. Сумма баллов за все вопросы равна 100.
- Проходной балл для зачета практического задания – 40 баллов.
- Работа должна быть сдана до окончания курса.

### Ситуация

Крупная авиакомпания разрабатывает новую стратегию и для этого изучает, кто является ее пассажирами, что они предпочитают, и от каких факторов зависит спрос на билеты. Была взята выборка 200 пассажиров. В опросе участвовали пассажиры, согласившиеся заполнить анкету. Каждый пассажир отвечал на вопросы о том,

- сколько раз он летал рейсами этой компании за последний год;
- какова цель его полетов (бизнес или личные цели);
- что он ценит больше: комфорт, точность или частоту рейсов (только один ответ);

также пассажира просили указать его

- пол,
- возраст,
- годовой доход.

Полученные данные сведены в таблицу, фрагмент которой вы видите

число полетов за последний год	пол	цель полета	самая важная характеристика	годовой доход, тыс.\$	возраст
1	женщина	бизнес	комфорт	22	35
3	мужчина	бизнес	комфорт	51	35
2	мужчина	бизнес	частота рейсов	18	42
2	мужчина	личные цели	точность	22	25
3	мужчина	личные цели	частота рейсов	36	28
3	мужчина	личные цели	точность	31	34

Скачав файл данных, Вы выполняете задания и отвечаете на вопросы.

Возможностей программы Excel достаточно для выполнения всех вычислений. Для некоторых заданий в Excel нет подходящих встроенных функций, и Вам придется провести расчет по формуле.

Вы можете также использовать другие статистические пакеты, которыми владеете (укажите какие именно программы).

Обратите внимание, что, выполняя задание, важно не просто получить некоторое значение или построить график, необходимо пояснить, что означает полученный результат.

Результаты выполнения всех заданий Вы оформляете в виде отчета.

## Вопросы и задания

1. Какие переменные в исследовании пассажиров являются количественными, а какие неколичественными – категориальными? (1 балл)

2. Постройте три сводные таблицы. (3 балла)

Указание. Воспользуйтесь меню-данные-сводная таблица.

Пол и цель полета

	цель полета		
пол	бизнес	личные цели	Общий итог
женщина			
мужчина			
Общий итог			

Пол и самая важная характеристика

	самая важная характеристика			
пол	комфорт	частота рейсов	точность	Общий итог
женщина				
мужчина				
Общий итог				

Цель полета и самая важная характеристика

	самая важная характеристика			
цель полета	комфорт	частота рейсов	точность	Общий итог
бизнес				
личные цели				
Общий итог				

3. Постройте две частотные таблицы: для распределения пассажиров по возрасту и по годовому доходу. (5 баллов)

Указание. Воспользуйтесь функцией «частота». Вставка-функция-статистические-частота. Прочитайте справку по функции. Подготовьте заранее массив интервалов (правые границы). Ячейки должны быть выделены одновременно, ОК нажимайте, удерживая Shift и Ctrl

Распределение пассажиров по возрасту

возраст		
левая граница интервала	правая граница интервала	частота
11	15	
16	20	
21	25	
26	30	
31	35	
36	40	
41	45	
46	50	
51	55	
56	60	
61	65	
66	70	
71	75	
всего		

Распределение пассажиров по годовому доходу

годовой доход		
левая граница интервала	правая граница интервала	частота
1	10	
11	20	
21	30	
31	40	
41	50	
51	60	
61	70	
71	80	
81	90	
91	100	
101	110	
111	120	
всего		

4. Представьте данные таблиц из п. 2 в графической форме. Выберите уместный тип диаграммы и поясните ее особенности.

В таблицы п. 3 добавьте столбики «средняя точка интервала» и «относительная частота» и постройте гистограммы. (8 баллов)

5. По данным таблицы из п.3 «Распределение пассажиров по годовому доходу» постройте кумулятивный частотный график («стрелку»). (3 балла)

6. Оцените метод, которым была сделана выборка. Какие возможны ошибки? Предложите подходящую технику выборки респондентов для этого исследования. Что может стать источником ошибок в вашем методе? (6 баллов)

7. По форме гистограмм из п.4 оцените, симметричными или несимметричными являются распределения пассажиров по возрасту и доходу.

Опишите данные по возрасту и доходу, используя уместные меры среднего и вариации. (5 баллов)

8. Определите 95% доверительные интервалы среднего числа полетов для тех пассажиров, кто летает с целью бизнеса, и тех, кто летает в личных целях. Поясните результаты. Указание. Проведите сортировку данных по столбцу «цель полета». (5 баллов)

9. Определите 95% доверительный интервал для доли женщин среди пассажиров авиакомпании. (5 баллов)

10. Для пассажиров-женщин определите 95% доверительный интервал доли тех, кто летает с личными целями. (4 балла)

11. Есть ли основания полагать, что среднее число полетов у мужчин больше, чем у женщин? (5 баллов)

12. Есть ли основания полагать, что среднее число полетов различно для тех пассажиров, кто летает с целью бизнеса, и тех, кто летает в личных целях? (5 баллов)

13. Существует ли связь между полом пассажира и целью полета? (3 балла)

14. Существует ли связь между полом пассажира и самой важной для него характеристикой полетов? (3 балла)

15. Существует ли связь между целью полета пассажира и наиболее важной для него характеристикой полета? (3 балла)

16. Есть ли основания полагать, что доля тех, кто ценит комфорт, больше для женщин? (5 баллов)

17. Изложите смысл результатов, полученных в п.п.11-16. (4 балла)

18. Постройте диаграммы рассеяния, чтобы показать связь между числом полетов и возрастом пассажира и числом полетов и доходом пассажира. Прокомментируйте форму. (3 балла)

19. Вычислите коэффициент корреляции между

- числом полетов и возрастом пассажира
- числом полетов и доходом пассажира
- возрастом пассажира и доходом пассажира

Для значимых коэффициентов корреляции вычислите коэффициент детерминации и поясните его смысл. (6 баллов)

20. Выберите наиболее приемлемую независимую переменную и постройте простую линейную регрессию. Проверьте значимость коэффициентов. (5 баллов)

21. Какие предположения сделаны при построении регрессии? Оцените эти предположения, используя график остатков. Можно ли использовать полученную модель как инструмент прогнозирования? (4 балла)

22. Спрогнозируйте число полетов пассажира со следующими характеристиками: женщина, летает с целью бизнеса, ценит точность, ей 45, она зарабатывает 32. Что влияет на вашу уверенность в прогнозе? (3 балла)

23. Учитывая результаты исследования, какие рекомендации Вы могли бы дать авиакомпании? (6 баллов)