

# Руководство пользователя

## Главная страница

Analyze

Histograms

Box-And-Whiskers

Scatter

Bar

Statistic

File name // Separator

insurance.csv // ,

File path

Select file

Separator

,

Загрузить файл

При нажатии на Analyze пользователь попадает на главную страницу сайта.

В боковом меню необходимо выбрать файл с исходными данными в формате и указать разделитель (для файла insurance.csv разделитель ‘,’ (запятая)) после чего необходимо нажать на кнопку «Загрузить файл». Индикацией успешной загрузки файла будет input поле «File name // Separator», где отобразится имя исходного файла и введенный разделитель.

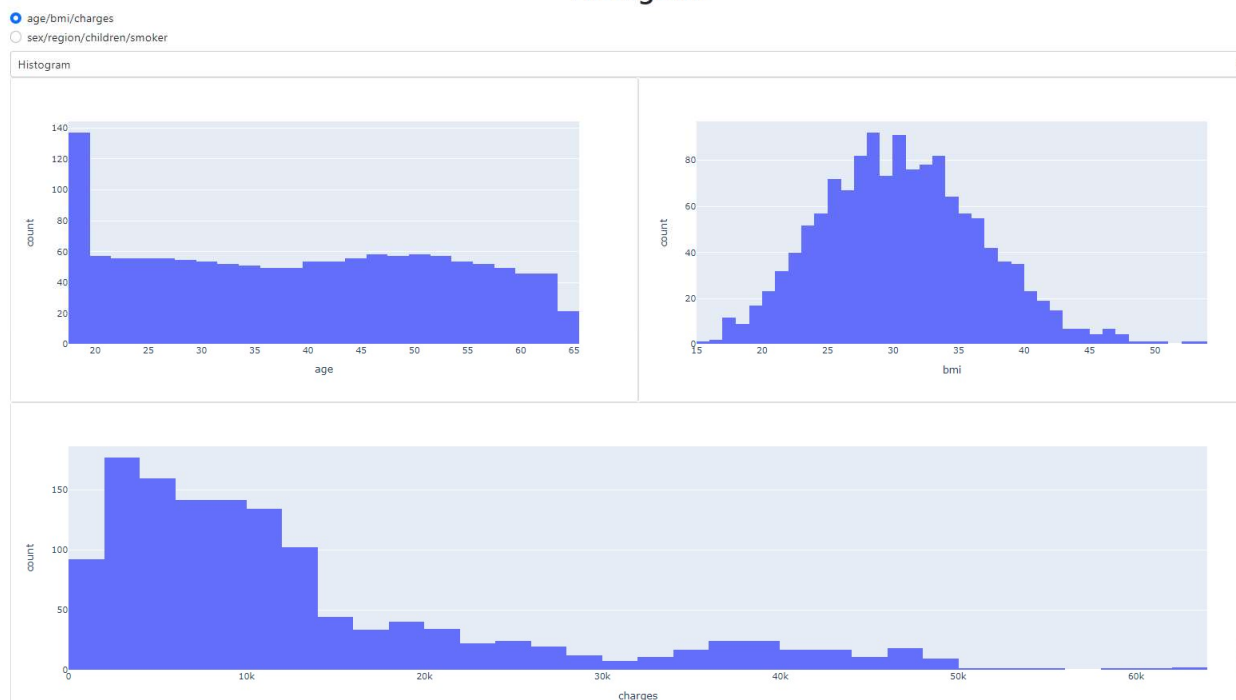
Рекомендуется использовать insurance.csv или insurance\_2.csv из директории приложения /assets (за такой короткий срок было проблематично сделать приложение максимально гибким, поэтому в некоторых местах присутствует hardcode на имена столбцов файла insurance.csv).

Особенности:

- 1) Файл insurance.csv может необязательно лежать именно в папке /assets – он может находиться в любом месте, так как при загрузке файла приложению отдается не путь до файла, а сразу сам файл в base64 (в виду этого есть проблемы с чтением файла в формате xls).
- 2) Контент других страниц не будет отображен, если не был загружен файл
- 3) Была реализована индикация «loader’a» рядом с названием страницы, чтобы пользователь понимал, что выбранное действие обрабатывается приложением

## Гистограммы

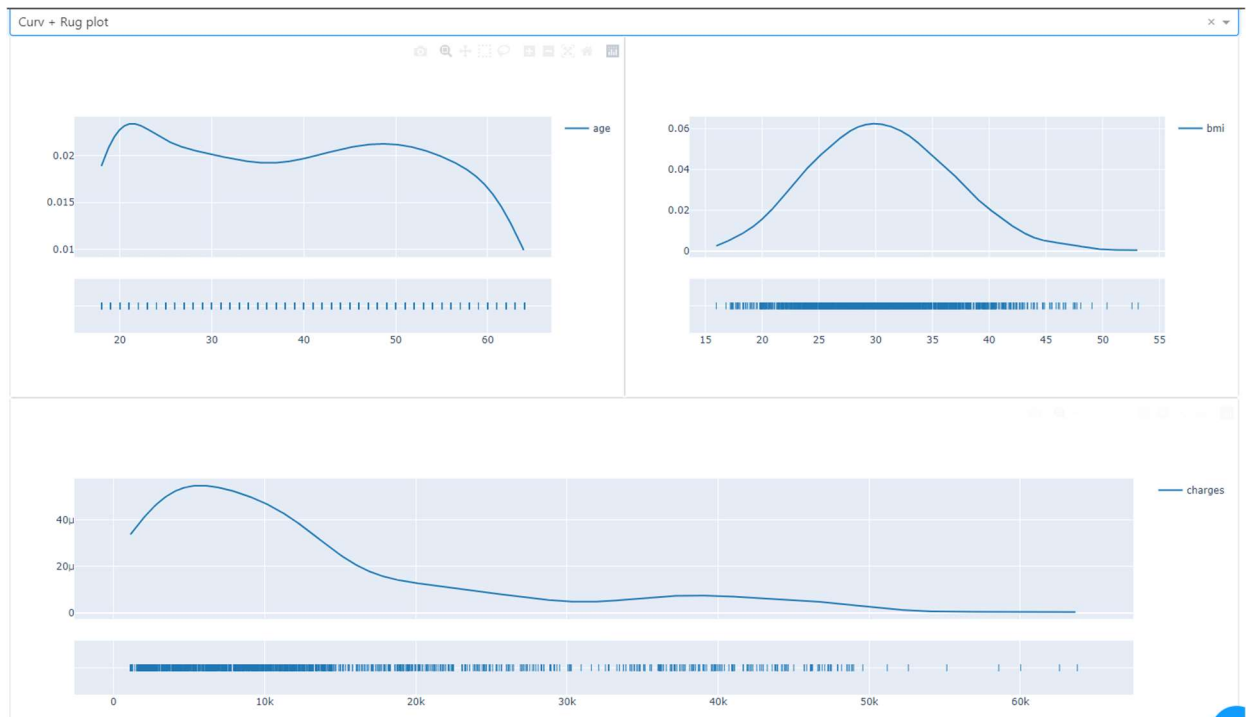
### Histogram



На странице Histogram отображаются гистограммы по параметрам:

- age/bmi/charges
  - возраст
  - индекс массы тела
  - цена мед. страховки
- sex/region/children/smoker
  - пол
  - регион
  - количество детей
  - курильщики

Для гистограммы age/bmi/charges можно выбрать Curv + Rug plot.



## Box-And-Whiskers diagram

Возможен выбор параметров для отображения по осям X и Y, а также группировка отображаемых параметров

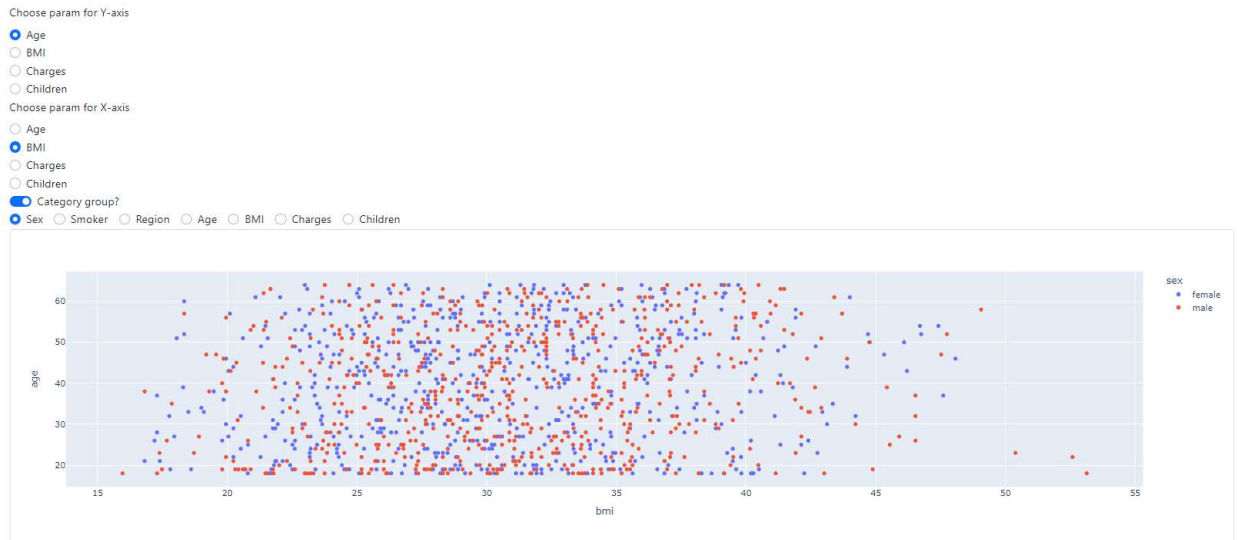
## Box-And-Whiskers diagram



## Scatter

Возможен выбор параметров для отображения по осям X и Y, а также группировка отображаемых параметров

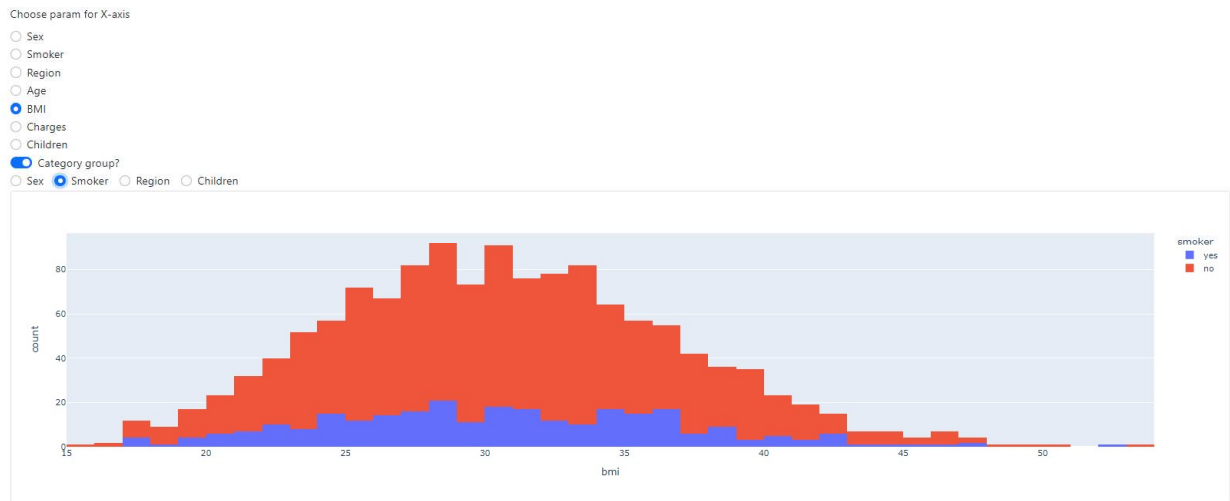
## Scatter



## Bar

Возможен выбор параметров для отображения по оси X, а также группировка отображаемых параметров

## Bar



## Statistic

Эта страница для анализа данных в табличном представлении.

### Statistic

☒ Show table?

☐ Sex ☐ Smoker ☒ Region ☒ Age ☐ BMI ☒ Charges ☐ Children

Age  min: 18, max: 64

AND

Charges  min: 1121.8739, max: 63770.42801

OR

Region

Number of lines  
  
total: 1130  
Current page  
  
last: 162

id	age	charges	region
0	19	16884.924	southwest
2	28	4449.462	southeast
3	33	21984.47061	northwest
4	32	3866.8552	northwest
5	31	3756.6216	southeast
6	46	8240.5896	southeast
7	37	7281.5056	northwest

☐ Show statistic table?

В «оранжевой» области есть возможность задать логическое условие для отбора нужных строк.

В «голубой» области есть возможность выбрать количество отображаемых строк и переключать «страницы» таблицы

Таблица генерируется при 3 взаимодействиях: при включении/выключении отображения столбцов (на картинке включены Region, Age, Charges), при взаимодействии с «голубой» областью и при нажатии кнопки «Update».

По кнопке «Clear» очищаются введенные логические условия

Особенности:

- 1) Значения, которые вводятся руками – валидируются, в случае некорректных значений окошко будет подсвечиваться красным
- 2) Для категориальных переменных нет input окошка, варианты, которые можно использовать генерируются автоматически в выпадающем списке при выборе названия категориальной переменной
- 3) Логика в приложении описана таким образом, что написанное логическое выражение не будет обрабатываться, если не указаны какие-либо поля. Также, если будет указано 3 логических условия, и они не будут объединены логическими операторами AND или OR, то

приложение посчитает, что было введено лишь самое верхнее логическое условие

- 4) Количество строк и столбцов высчитываются автоматически каждый раз при изменении таблицы, поэтому в голубой области всегда будут подсказки по количеству строк, которые можно отобразить на текущей странице и количество страниц в целом.
- 5) Похожие подсказки появляются у числовых полей в логическом условии, отображается минимальное и максимальное возможное значение, для лучшего ориентирования пользователя в данных

## Statistic

☐ Show table?

☒ Show statistic table?

☒ Age ☒ BMI ☒ Charges

Имя	Среднее	Медиана	Станд. отклонение	Межквартильный размах	Верхняя квартиль	Нижняя квартиль	Коэффициент асимметрии	Коэффициент эксцесса	Количество наблюдений	Количество пропущенных значений
Age	39.20702541106129	39	14.049960379216154	24	51	27	0.05567251565299186	-1.2450876526418673	1338	0
Bmi	30.66339686098655	30.4	6.098186911679014	8.3975	34.69375	26.29625	0.2840471105987448	-0.05073153135467834	1338	0
Charges	13270.422265141257	9382.033	12110.011236694001	11899.625365	16639.912515	4740.28715	1.5158796580240388	1.6062986532967907	1338	0

В этой таблице можно посмотреть статистическую информацию по числовым полям (не категориальным).

Таблица генерируется при выборе необходимых столбцов (на картинке выбраны Age, BMI, Charges)