Отчет учащегося 3 курса

Факультете компьютерных технологий и

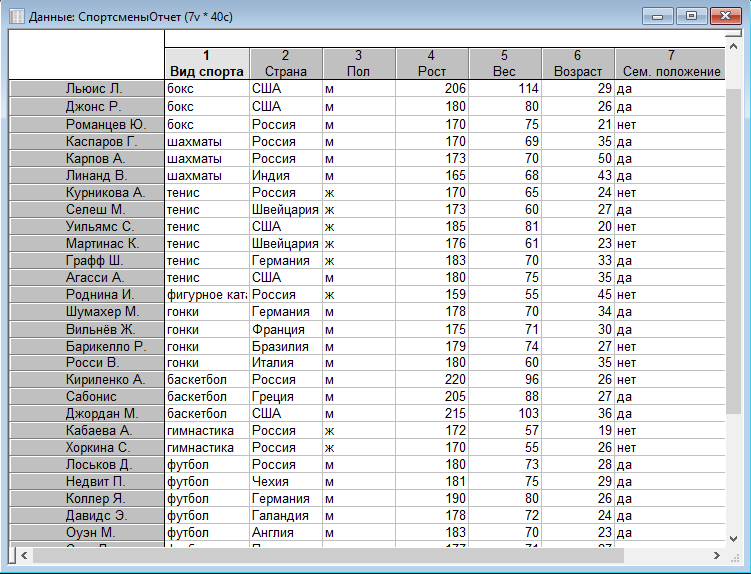
Прикладной Математики

Студента 37 группы

Павлова Дмитрия Павловича

Краснодар 2023

В данной работе была использована следующих таблиц сходных данных:



В которой имеются следующие переменные:

Вид спорта – категориальная переменная , указывающая каким видом спорта занимается спортсмен.

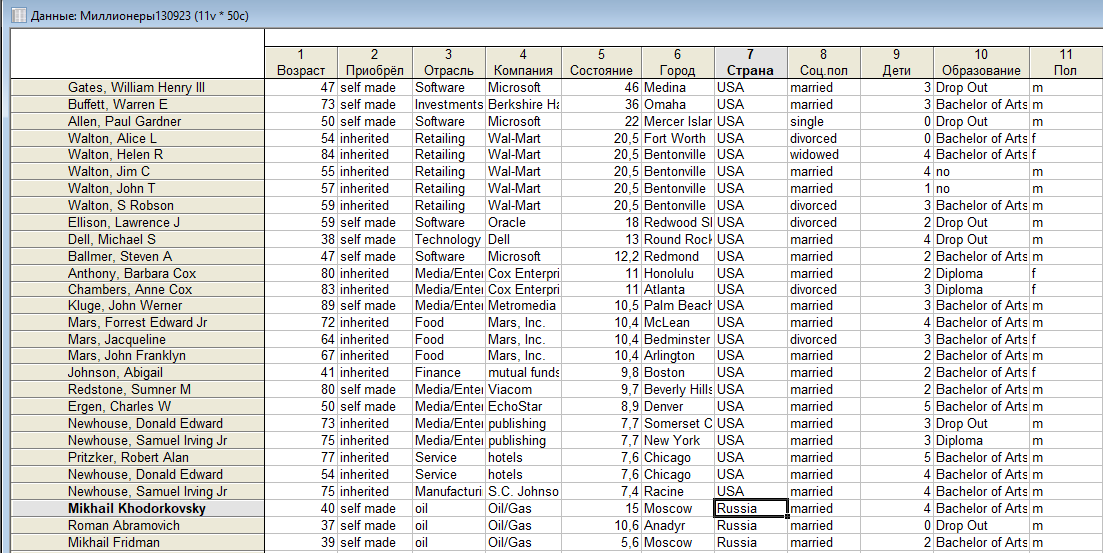
Страна – категориальная переменная , указывающая к какой стране относится спортсмен .

Пол- категориальная переменная , указывающая пол спортсмена

Рост – непрерывная , числовая переменная , указывающая рост спортсмена .

Вес – числовая переменная , указывающая вес спортсмена

Семейное положение – категориальная переменная , характеризующая семейное положение .



В данной таблице исходных данных имеются следующие переменные:

Возраст – числовая переменная , указывающая возраст миллионера на момент измерения

Приобрел – категориальная переменная , указывающая как была получена компания (создал сам либо уначледовал).

Отрасль – категориальная переменная , указывающая отрасль деятельности .

Компания – категориальная переменная , указывающая компанию с которой связан миллионер .

Состояния – числовая переменная , указывающая состояние миллионера в миллионах долларов .

Город – текстовая переменная , указывающая город происхождения миллионера .

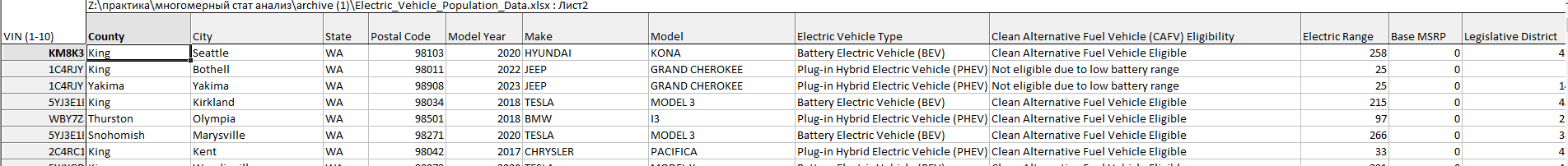
Страна – категориальная переменная , указывающая из какой страны миллионер .

Социальное положение – категориальная переменная , указывающая семейное положение .

Дети – числовая переменная , указывающая количество детей

Образование – категориальная переменная , указывающая образование миллионера

Пол – категориальная переменная , указывающая пол миллионера



Изображение выглядит как текст, Шрифт, число, линия

Автоматически созданное описание

В данной таблице исходных данных имеются следующие переменные:

Country – категориальная переменная , указывающая страну автомобиля

City – текстовая переменная , указывающая город автомобиля

State – категориальная переменная , указывающая состояние автомобиля

postal code – числовая переменная , указывающая почтовый индекс автомобиля

Model Year – числовая переменная , указывающая год выпуска автомобиля

Make – категориальная переменная , указывающая производителя автомобиля

Model – текстовая переменная , указывающая модель автомобиля

Electric Vehicle Type – категориальная переменная , указывающая тип электромобиля

Clean Alternative Fuel Vehicle (CAFV) Eligibility – категориальная переменная , указывающая право на использование экологически чистых транспортных средств на альтернативном топливе (CAFV)

Electric Range – числовая переменная , указывающая электрический диапазон

Base MSRP – числовая переменная , указывающая рекомендуемую розничную стоимость

Legislative District – rчисловая переменная указывающая законодательный округ

DOL Vehicle ID – числовая переменная , указывающая DOL идентификатор автомобиля

Vehicle Location – непрерывная переменная , указывающая местоположение автомобиля

Electric Utility – категориальная переменная , указывающая тип электроснабжения автомобиля

2020 Census Tract – числовая переменная , указывающая место в рейтинге в 2020 году

Изображение выглядит как текст, снимок экрана, программное обеспечение, число

Автоматически созданное описание

На данном слайде показана таблица корреляций между переменными таблицы спортсмены .

Чем ближе коэффициент корреляции к единице по модулю , тем сильнее корреляция между переменными .

Красным цветом в ячейках таблицы отмечены статистически значимые коэффициенты корреляции .

Из таблицы видно , что наибольшая взаимосвязь имеется между переменными рост и вес . Так как коэффициент корреляции у этих переменных больше нуля , увеличение или уменьшение значения одной переменной , приводит к увеличению (уменьшению ) значения другой . В случае если коэффициент корреляции меньше нуля, увеличение (уменьшение ) значения одной , приводит к уменьшению (увеличению) значения другой .

Наибольший по модулю , но отрицательный по знаку коэффициент корреляции имеют переменные Вес и Пол .

Если коэффициент корреляции по модулю близок к нулю , то взаимосвязь между переменными слабая , либо вообще отсутствует . Такая ситуация наблюдается между переменными Вид спорта и семейное положение .

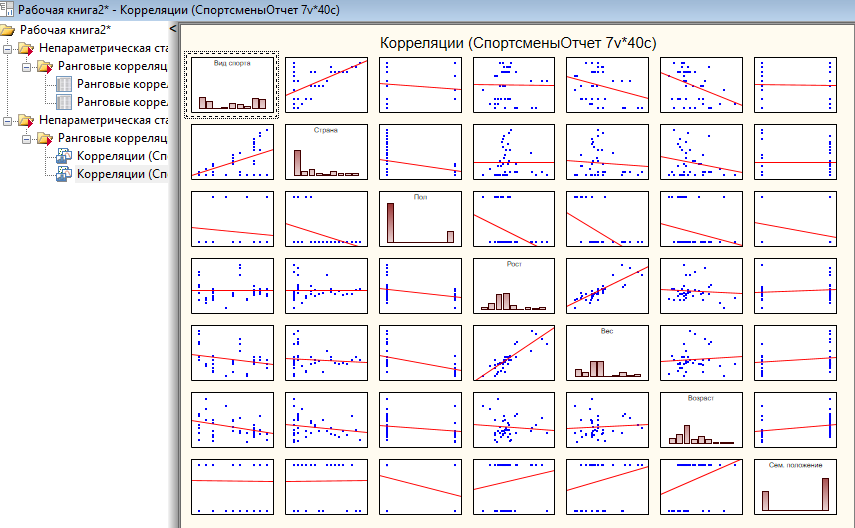
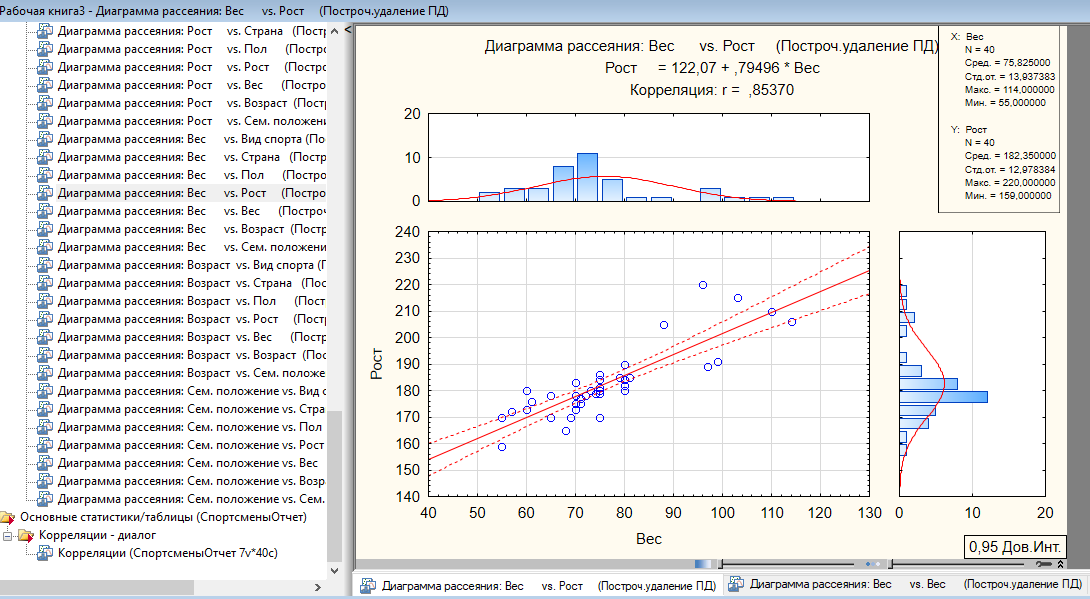


График корреляций между переменными

На картинках отображаются графики корреляций между переменными . Чем ближе синие точки к красной корреляционной линии , тем сильнее взаимосвязь . Если корреляционная линия убывает , то коэффициент корреляции меньше нуля , если возрастает , то больше .



На слайде показана диаграмма рассеивания для переменных Вес и Рост .

Верхний график отображает соотношение прогнозируемого (красная линия ) и фактического (Синие гистограммы ) значений для переменной Вес . Аналогичные диаграммы строятся для переменной Рост (Расположен под таблицей ) .

График расположенный между диаграммами , отображает коэффициент корреляции между переменными Рост и вес . Также , на графике присутствуют пунктирные линии , отображающие доверительный интервал , отклонение значений от корреляционной прямой в пределах которого является допустимым .

Над диаграммой веса , указывается коэффициент корреляции r , а над ним , формула взаимосвязи значений переменных рост и вес . В таблице приведены дополнительные сведения для данных переменных .

Так , это общее количество записей в таблице для каждой из переменных , среднее значение для каждой из них . Также , это величина стандартного отклонения , а также минимальное и максимальное значение .