

STAT394 Group Project Report1 TEST

Ken MacIver, Tom Tribe, Jundi Yang, Mei Huang

2022-08-08

Contents

1	First Project Report TEST	1
1.1	Introduction	1
1.2	Platform and Data	1
1.3	Exploratory Data Analysis (EDA)	2

Load ggplot and ggthemes

1 First Project Report TEST

1.1 Introduction

We report on the dataset we have selected for our Group Project. It contains data on athletes from different sports. There are two categorical variables, ‘Gender’ and ‘Sport’, and seven continuous variables which are: weight, height, body mass index (BMI) etc, etc.

1.2 Platform and Data

The data comes from the website of the Australasian Data and Story Library (OzDASL) available at the following link <http://www.statsci.org/data/oz/ais.txt>, who have kindly allowed their data to be used for statistical teaching and learning purposes.

Categorical data:

- Sex = male or female
- Sport = sport played

Numeric data:

- Ht = height in cm
- Wt = weight in kg
- LBM = lean body mass in kg
- %Bfat = % body fat
- BMI = body mass index (weight/height²)

- RCC = red blood cell count
- WCC = white blood cell count
- Hc = haematocrit
- Hg = haemoglobin
- Ferr = plasma ferritin concentration
- SSF = sum of skin folds

1.3 Exploratory Data Analysis (EDA)

In this section we perform an exploratory analysis to see what sort of patterns emerge from the data.

Firstly, we load the data into R:

```
# read data into R and display the first ten columns
# use relative file path to allow other users to load the dataset
sports <- read.csv("./Sports_data2.csv")

sports[1:10,]
```

##	i..Sex	Sport	RCC	WCC	Hc	Hg	Ferr	BMI	SSF	X.Bfat	LBM	Ht	Wt
## 1	female	BBall	3.96	7.5	37.5	12.3	60	20.56	109.1	19.75	63.32	195.9	78.9
## 2	female	BBall	4.41	8.3	38.2	12.7	68	20.67	102.8	21.30	58.55	189.7	74.4
## 3	female	BBall	4.14	5.0	36.4	11.6	21	21.86	104.6	19.88	55.36	177.8	69.1
## 4	female	BBall	4.11	5.3	37.3	12.6	69	21.88	126.4	23.66	57.18	185.0	74.9
## 5	female	BBall	4.45	6.8	41.5	14.0	29	18.96	80.3	17.64	53.20	184.6	64.6
## 6	female	BBall	4.10	4.4	37.4	12.5	42	21.04	75.2	15.58	53.77	174.0	63.7
## 7	female	BBall	4.31	5.3	39.6	12.8	73	21.69	87.2	19.99	60.17	186.2	75.2
## 8	female	BBall	4.42	5.7	39.9	13.2	44	20.62	97.9	22.43	48.33	173.8	62.3
## 9	female	BBall	4.30	8.9	41.1	13.5	41	22.64	75.1	17.95	54.57	171.4	66.5
## 10	female	BBall	4.51	4.4	41.6	12.7	44	19.44	65.1	15.07	53.42	179.9	62.9

Display scatterplots of all the variables against each other.

```
plot(sports)
```

