Assignment 1

Simon Clarke David Ewing Xia Yu

2025-03-09

Braking Distance

In this question, do not use the `lm` function or a module that provides an implementation of k-NN. You are allowed to use elementary statistical objects like mean, variance, etc.

We will be predicting the distance that a car takes to stop when driving at a certain speed. The dataset is from 1930, so it might be slightly outdated. Units are miles per hour (speed) and feet (distance).

Data Preparation

```
# Load and preprocess dataset
```

Linear Regression (Without lm)

```
# Compute slope and intercept for simple linear regression
```

Using the linear regression model, predict the braking distance for a car going at 30 km/h and include an 80% prediction interval.

```
# Prediction for 30 km/h
```

k-NN Model

```
# Fit and predict using k-NN model
```

Filipino Household Income

Data Preparation

```
# Load and preprocess dataset
```

Linear Regression

```
# Fit linear model and summarise
```

Predicting Possum Age

Data Preparation

```
# Load and preprocess dataset
 zip_path <- "../data/datasets.zip"
 file.exists(zip_path)
```

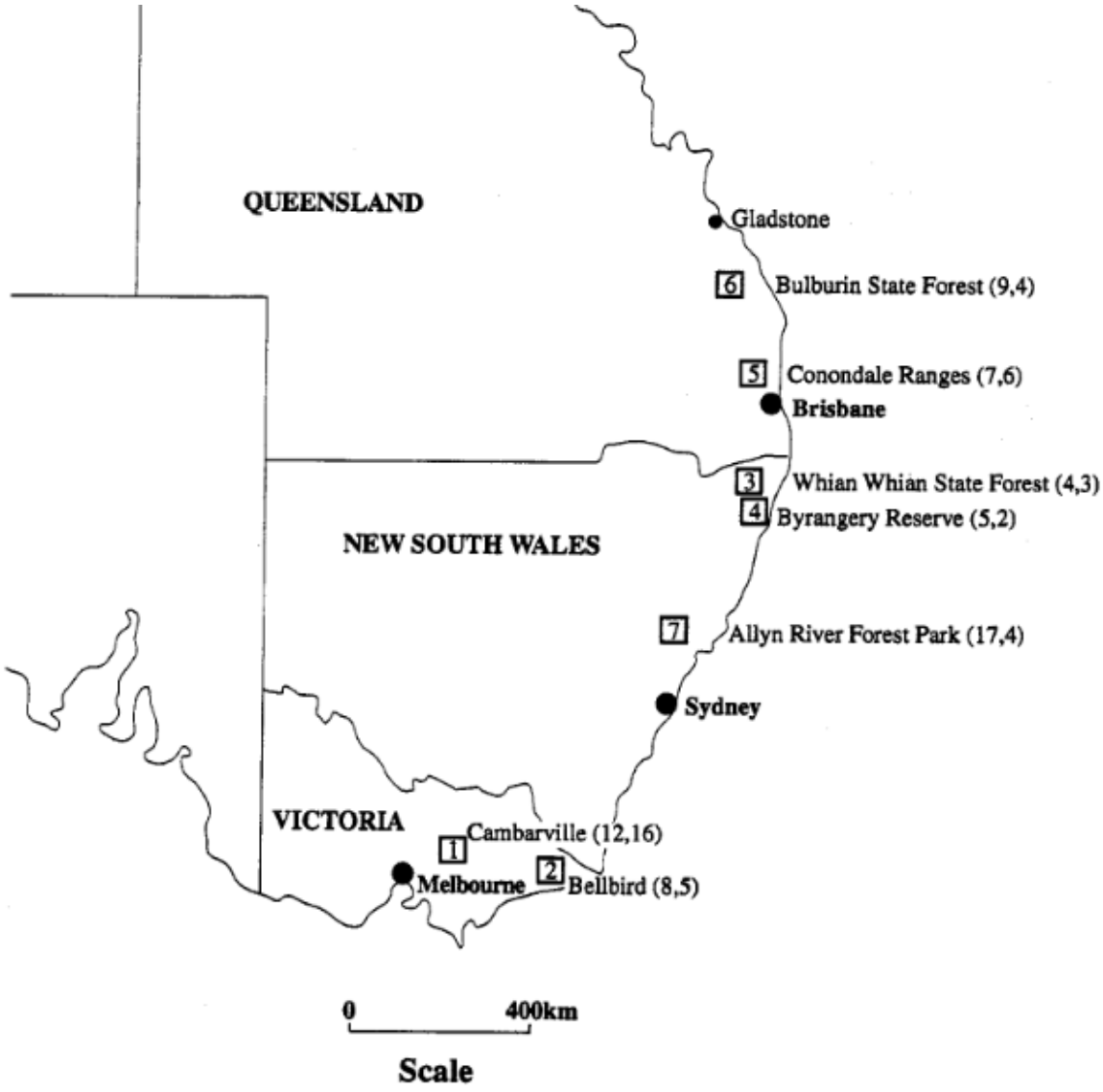
```
## [1] TRUE
```

```
 possums <- read.csv(unz(zip_path, "possums.csv"))
 head(possums)
```

```
##   case site Pop sex age hdlngth skullw totlngth taill footlgth earconch  eye
## 1    1    1 Vic   m   8    94.1   60.4     89.0  36.0     74.5     54.5 15.2
## 2    2    1 Vic   f   6    92.5   57.6     91.5  36.5     72.5     51.2 16.0
## 3    3    1 Vic   f   6    94.0   60.0     95.5  39.0     75.4     51.9 15.5
## 4    4    1 Vic   f   6    93.2   57.1     92.0  38.0     76.1     52.2 15.2
## 5    5    1 Vic   f   2    91.5   56.3     85.5  36.0     71.0     53.2 15.1
## 6    6    1 Vic   f   1    93.1   54.8     90.5  35.5     73.2     53.6 14.2
##   chest belly
## 1  28.0    36
## 2  28.5    33
## 3  30.0    34
## 4  28.0    34
## 5  28.5    33
## 6  30.0    32
```

Data and Initial Analysis

```
# Load dataset and visualise
knitr::include_graphics("../images/possum_age_plot.png")
```




```
## Data Preparation
```

```
# Preprocess dataset
```

Feature Selection and Model Training

```
# Forward feature selection and model training
```

Model Evaluation

```
# Compute evaluation metrics
```

Further Exploration

```
# Additional analysis or research questions
```