

Assessment 1 - 01: Data and Introduction

The brief

Talk about setting up the idea of being an insurance firm and wanting to be able to predict whether 'client' are smokers or not from data we have collected to then assign a higher premium.

The team

We are a four person team consisting of, *Jiaqi*, *Revati*, *Yifan*, and *Hugh*.

Library requirements

! Lib requiremnts should be provided further in individual analyses

We need `fs` to be able to handle directories in a cross-platform way. This makes the whole analysis much more repeatable and is good data science practice.

```
if (!require("fs")) install.packages("fs")
```

```
## Loading required package: fs
```

```
library("fs")
```

The data

Get the data

First create a directory for it to go in:

```
rawdatadir=path_wd("../data", "raw")  
if(!dir.exists(rawdatadir)) dir.create(rawdatadir,recursive = TRUE)
```

Download it

```
download.file("https://raw.githubusercontent.com/EurousZhao/DST-group_A1/main/Datasets/smokin  
g_drinking_dataset.csv",path_wd("../data", "raw", "smoking_drinking_dataset.csv"))
```

```
data <- read.table(path_wd("../data", "raw", "smoking_drinking_dataset.csv"),header=TRUE, sep  
= ",")
```