

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH



UIT
Trường Đại học
Công nghệ Thông tin

**Khoa Khoa học
và Kỹ thuật Thông tin**

CODEBOOK MÔ TẢ BỘ DỮ LIỆU BANK MARKETING

Môn học: Thu thập và tiền xử lý dữ liệu - DS103.N21

Tên: Nguyễn Thị Huyền Trang

MSSV: 21520488

TP HỒ CHÍ MINH, 2023

MỤC LỤC

CODEBOOK	3
Bộ dữ liệu	3
Codebook	3
Raw data	4
Tidy data	4
Instruction list	4
THAM KHẢO	5

CODEBOOK

Bộ dữ liệu: Bank Marketing

Codebook:

Thông tin	Nội dung
Tên bộ dữ liệu	Bank Marketing
Nguồn thu thập và cách thức thu thập	<p>Dữ liệu có liên quan đến các chiến dịch tiếp thị trực tiếp của một ngân hàng ở Bồ Đào Nha.</p> <p>Các chiến dịch tiếp thị được thực hiện dựa trên các cuộc gọi điện thoại. Thông thường, sẽ cần có nhiều hơn một địa chỉ liên hệ với cùng một khách hàng để xem sản phẩm (tiền gửi có kỳ hạn ngân hàng) sẽ được đăng ký (hoặc không).</p>
Số thuộc tính	17
Kích thước	45211
Thông tin tên các thuộc tính	<p>Age: Độ tuổi</p> <p>Job: Công việc</p> <p>Marital: Tình trạng hôn nhân</p> <p>Education: Mức độ hoàn thành chương trình giáo dục</p> <p>Default: Có tính dụng trong tình trạng vỡ nợ</p> <p>Balance: Số dư trung bình hằng năm</p> <p>Housing: Có căn nhà nào đang được thế chấp</p> <p>Loan: Khách hàng có khoảng vay cá nhân không</p> <p>Contact: Cách thức liên lạc</p> <p>Day: Ngày liên hệ cuối cùng của tháng</p> <p>Month: Tháng liên hệ cuối cùng trong năm</p> <p>Duration: Thời lượng lần cuối cùng liên lạc</p> <p>Campaign: Số liên hệ được thực hiện trong chiến dịch này và cho khách hàng này</p> <p>Pdays: số ngày trôi qua sau khi khách hàng được liên hệ lần cuối từ một chiến dịch trước đó. Trong đó: -1 có nghĩa là khách hàng này chưa được liên hệ trước đó</p> <p>Previous: Số lượng địa chỉ liên hệ được thực hiện trước chiến dịch này và cho khách hàng này</p>

	Poutcome: Kết quả của chiến dịch tiếp thị trước đó Y: Khách hàng đã đăng ký tiền gửi có kỳ hạn chưa?
Thông tin tác giả	Paulo Cortez (Univ. Minho) và Sérgio Moro (ISCTE-IUL) @ 2012

Raw data:

Raw data gồm file: Bank-full.csv (chứa tất cả dữ liệu).

Tidy data:

Xử lý file Bank-full.csv, dữ liệu sau khi xử lý được lưu dưới file bank.csv

Instruction list:

```
rm(list=ls())
myFiles <- list.files(path="bank-data/", pattern="bank-full.csv")
file <- read.csv(paste("bank-data/", myFiles, sep=""), sep=";", header = FALSE)
variables <-file[c(1),]
colnames(file)<-variables
dataset<-file[-c(1),]
row.names(dataset) <- NULL
write.csv(dataset, file = "bank.csv",row.names = FALSE)
```

THAM KHẢO

Bank Marketing Data Set, [UCI Machine Learning Repository: Bank Marketing Data Set](#)