**Phân tích nhân khẩu học khách hàng cá nhân của một ngân hàng**

1. **Dữ liệu đầu vào**

* Các file dữ liệu aum.csv, prod\_holding.csv, cust.csv

|  |  |
| --- | --- |
| File **aum.csv**: thông tin tổng tài sản mà khách hàng nắm giữ | |
| customer\_id | Mã số định danh của khách hàng |
| amount | Tổng tài sản khách hàng nắm giữ |

|  |  |
| --- | --- |
| File **cust.csv**: thông tin về phân khúc và nơi ở của khách hàng | |
| customer\_id | Mã số định danh của khách hàng |
| segment | Phân khúc khách hàng |
| province\_city | Tỉnh/thành phố nơi khách hàng sinh sống |

|  |  |
| --- | --- |
| File **prod\_holding.csv**: thông tin về các sản phẩm mà khách hàng sử dụng | |
| customer\_id | Mã số định danh của khách hàng |
| prod\_ca | Sản phẩm tài khoản thanh toán |
| prod\_td | Sản phẩm tiền gửi có kỳ hạn |
| prod\_credit\_card | Sản phẩm thẻ tín dụng |

1. **Mục tiêu**

* Phân tích tổng quan khách hàng theo từng phân khúc (Regular, Silver, Gold)

1. **Các yêu cầu, câu hỏi và kết quả mong muốn**

* Import dữ liệu từ các file csv và thực hiện các thao tác kiểm tra, tìm hiểu bộ dữ liệu
* Đổi tên cột customer\_id thành cust\_id
* Số lượng và Tỷ trọng khách hàng theo từng phân khúc là bao nhiêu? Tạo bảng gồm các cột như dưới đây sau đó vẽ bar chart thể hiện số lượng khách hàng

|  |  |  |
| --- | --- | --- |
| Segment | Number Of Customers | Proportion (%) |

* Khách hàng của từng phân khúc tập trung chủ yếu ở những tỉnh, thành phố nào? Sắp xếp theo thứ tự từ lớn nhất đến nhỏ nhất của các cột Regular, Silver, Gold.

|  |  |  |  |
| --- | --- | --- | --- |
| Province/City | Regular | Silver | Gold |

* AUM của khách hàng (Min, Q25, Q50, Q75, Q90, Max, Mean) giữa các phân khúc là như thế nào? (tạo bảng và vẽ box plot cho 3 segment)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Segment | Min | Q25 | Q50 | Q75 | Q90 | Max | Mean |

* Tỷ lệ nắm giữ các sản phẩm (Product Penetration) Credit Card, Current Account, Term Deposit của tập khách hàng Regular là như thế nào?

|  |  |  |  |
| --- | --- | --- | --- |
| Product | No of customers using product | No of customers not using product | Product penetration rate (tỷ lệ khách hàng sở hữu sản phẩm) |