**BỘ MÔN HỆ THỐNG THÔNG TIN – KHOA CÔNG NGHỆ THÔNG TIN**

**ĐẠI HỌC KHOA HỌC TỰ NHIÊN THÀNH PHỐ HỒ CHÍ MINH, ĐẠI HỌC QUỐC GIA TP HCM**

**BẢNG THÔNG TIN CHI TIẾT NHÓM**

| **Mã nhóm:** | 07 | |
| --- | --- | --- |
| **Số lượng:** | 04 | |
| **MSSV** | **Họ tên** | **Email** |
| 20127361 | Nguyễn Ngọc Diệu Trang | 20127361@student.hcmus.edu.vn |
| 20127632 | Trần Đình Thiện | 20127632@student.hcmus.edu.vn |
| 21127620 | Trần Hoàng Khải | 21127620@student.hcmus.edu.vn |
| 21127705 | Từ Phước Toàn | 21127705@student.hcmus.edu.vn |

| **Bảng phân công & đánh giá hoàn thành công việc** | | | | |
| --- | --- | --- | --- | --- |
| **Công việc thực hiện** | **Người thực hiện** | **Phần trăm công việc** | **Mức độ hoàn thành** | **Đánh giá của nhóm** |
| Data abstraction & data profiling: invoice id, branch, customer type | Nguyễn Ngọc Diệu Trang | 25% | 100% | 10/10 |
| Task abstraction, idiom và vẽ biểu đồ: So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh; Thống kê tổng doanh thu bán hàng của các chi nhánh |
| Data abstraction & data profiling: gender, productid, quantity, tax 5% | Trần Hoàng Khải | 25% | 100% | 10/10 |
| Task abstraction, idiom và vẽ biểu đồ: Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?; So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020? |
| Data abstraction & data profiling: date, time, payment, cogs | Từ Phước Toàn | 25% | 100% | 10/10 |
| Task abstraction, idiom và vẽ biểu đồ: Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất; Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B |
| Data abstraction & data profiling: gross margin percentage, gross income, rating | Trần Đình Thiện | 25% | 100% | 10/10 |
| Task abstraction, idiom và vẽ biểu đồ: Sản phẩm nào bán chạy nhất ?, So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ? |

MỤC LỤC

[**I.**](#_heading=h.gjdgxs) **DATA ABSTRACTION VÀ DATA PROFILING** 4

[**1.**](#_heading=h.30j0zll) **Data abstraction** 4

[**2.**](#_heading=h.1fob9te) **Data profiling** 12

[**II.**](#_heading=h.3znysh7) **TASK ABSTRACTION** 12

[**1.**](#_heading=h.2et92p0) **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh** 12

[**2.**](#_heading=h.tyjcwt) **Thống kê tổng doanh thu bán hàng của các chi nhánh** 12

[**3.**](#_heading=h.3dy6vkm) **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất** 12

[**4.**](#_heading=h.1t3h5sf) **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B** 12

[**5.**](#_heading=h.4d34og8) **Sản phẩm nào bán chạy nhất ?** 12

[**6.**](#_heading=h.2s8eyo1) **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?** 12

[**7.**](#_heading=h.17dp8vu) **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?** 13

[**8.**](#_heading=h.3rdcrjn) **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?** 13

[**III.**](#_heading=h.26in1rg) **IDIOM** 13

[**1.**](#_heading=h.lnxbz9) **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh** 13

[**2.**](#_heading=h.35nkun2) **Thống kê tổng doanh thu bán hàng của các chi nhánh** 13

[**3.**](#_heading=h.1ksv4uv) **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất** 14

[**4.**](#_heading=h.44sinio) **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B** 14

[**5.**](#_heading=h.2jxsxqh) **Sản phẩm nào bán chạy nhất ?** 14

[**6.**](#_heading=h.z337ya) **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?** 14

[**7.**](#_heading=h.3j2qqm3) **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?** 15

[**8.**](#_heading=h.1y810tw) **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?** 15

[**IV.**](#_heading=h.4i7ojhp) **BIỂU ĐỒ VÀ NHẬN XÉT** 15

[**1.**](#_heading=h.2xcytpi) **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh** 16

[**2.**](#_heading=h.1ci93xb) **Thống kê tổng doanh thu bán hàng của các chi nhánh** 16

[**3.**](#_heading=h.3whwml4) **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất** 17

[**4.**](#_heading=h.2bn6wsx) **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B** 18

[**5.**](#_heading=h.qsh70q) **Sản phẩm nào bán chạy nhất ?** 19

[**6.**](#_heading=h.3as4poj) **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?** 19

[**7.**](#_heading=h.1pxezwc) **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?** 20

[**8.**](#_heading=h.49x2ik5) **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?** 21

[**Reference** 21](#_heading=h.2p2csry)

1. **DATA ABSTRACTION VÀ DATA PROFILING**
2. **Data abstraction**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a data sheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a document

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

A screenshot of a spreadsheet

Description automatically generated

1. **Data profiling**

A screenshot of a computer

Description automatically generated

1. **TASK ABSTRACTION**
2. **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh**

Produce -> browse -> compare

1. **Thống kê tổng doanh thu bán hàng của các chi nhánh**

Produce -> explore -> Sum

1. **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất**

Consume -> locate -> identify

1. **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B**

Produce -> browse -> Summarize

1. **Sản phẩm nào bán chạy nhất ?**

Produce -> explore -> identify

1. **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?**

Produce -> explore -> compare

1. **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?**

Produce -> locate -> compare

1. **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?**

Produce -> lookup -> Sum

1. **IDIOM**
2. **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh**

| Idiom | Stacked Bar Charts |
| --- | --- |
| Data Type | 2 categorical:  Customer type, branch  1 quantitative: totalInvoice |
| Encoding | Mark: bar mark (vertical)   * Glyph: multiple sub-bars are stacked vertically   Channel:   * Position: vertical and horizontal * Length, color HUE * Spatial region: glyph * Align: full glyph, lowest bar component * Separate: horizotal position |
| Task | Part-to-whole relationship |
| Scale | Stacked key attribute: 2 levels  Main key attribute: 3 levels |

1. **Thống kê tổng doanh thu bán hàng của các chi nhánh**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical:  Branch  1 quantitative: totalSales |
| Encoding | Mark: line  Channel:   * Position: vertical and horizontal * Size: length, area * Express: quantitative (length) * Spatial region: categorical * Separate: horizontal position * Align: vertical positionOrder: label (alphabeltic) * Color: HUE |
| Task | Compare, lookup |
| Scale | Key attribute: 3 levels |

1. **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: invoice ID  1quantitative: Total |
| Encoding | Mark: line  Channel:   * Position:Both * Length, color HUE * Size: length, Area * Spatial region: Categorical * Separate: horizotal position |
| Task | Find the highest bill |
| Scale | Top 10 items |

1. **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: invoice ID  1quantitative: Rating |
| Encoding | Mark:line (vertical)  Channel:   * Position: Both * Length, color HUE * Spatial region: Categorical |
| Task | Compare, lookup value |
| Scale | All items |

1. **Sản phẩm nào bán chạy nhất ?**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: ProductID  1 quantitative: Quantity |
| Encoding | Express: quantitative  Spatial region: categorical  Separate: vertical position  Align: horizontal position  Order:height |
| Task | Compare, lookup value |
| Scale | 943 items |

1. **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: YearWorked  1quantitative: Payment |
| Encoding | Channel:   * Position: horizontal and vertical * Align: lengths (quantitative) * Separate, order: horizontal region (ordered) |
| Task | Compare |
| Scale | 2 levels |

1. **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: Gender  1quantitative: Total |
| Encoding | Mark: line  Channel:   * Position:Both * Length, color HUE * Size: length, Area * Spatial region: Categorical * Separate: horizotal position |
| Task | Compare |
| Scale | All items |

1. **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?**

| Idiom | Bar Charts |
| --- | --- |
| Data Type | 1 categorical: Year  1quantitative: Gross income |
| Encoding | Mark: line  Channel:   * Position:Both * Length, color HUE * Size: length, Area * Spatial region: Categorical * Separate: horizotal position |
| Task | Compare |
| Scale | All items |

1. **BIỂU ĐỒ VÀ NHẬN XÉT**
2. **So sánh số lượng hóa đơn bán hàng giữa loại khách hàng ‘Normal’ và ‘Member’ theo từng chi nhánh**

A graph of blue and red squares

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

* **Đánh giá:**
* Nhìn vào biểu đồ trên ta thấy tổng số hóa đơn bán hàng của 2 loại khách hàng normal và member của chi nhánh A là lớn nhất
* Nhìn vào biểu đồ trên ta thấy tổng số hóa đơn bán hàng của 2 loại khách hàng normal và member của chi nhánh C là nhỏ nhất nhất
* Độ chênh lệch giữa tổng số hóa đơn của các chi nhánh là không quá lớn

1. **Thống kê tổng doanh thu bán hàng của các chi nhánh**

A graph with different colored bars

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

* **Đánh giá**
* Tổng doanh thu của chi nhánh C là cao nhất và chi nhánh B là thấp nhất
* Không có sự chênh lệch quá lớn giữa doanh thu của các chi nhánh

1. **Tìm hóa đơn của khách hàng có tổng tiền thanh toán ( Tính cả thuế) cao nhất**

A graph with blue lines

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá:** Biểu đồ trên thể hiện top 10 ngườ có hóa đơn thanh toán cao nhất được xếp thứ tự từ cao nhất đến thấp nhất, từ đó ta có thể biết hóa đơn người mua hàng 860-79-0874 với hóa đơn giá trị 1042,65 là cao nhất

1. **Tính đánh giá trung bình cho toàn bộ đơn hàng của chi nhánh B**

A red blue and yellow rectangular bars

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá**: Như Biều đồ ta thấy bình đánh giá ở chi nhánh B là 6,7914, một con số ổn định nhưng lại thấp nhất trong cả 3 chi nhánh của cửa hảng

1. **Sản phẩm nào bán chạy nhất ?**

A black rectangular object with text

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá**: Như Biều đồ ta thấy sản phẩm bán chạy nhất có ProductID là PID915 . Nhìn về các sản phẩm khác ta thấy các sản phẩm đều gần như số lượng mua như nhau và chỉ chênh nhau từ 1 đến gần 20 sản phẩm. Từ đó ta nhận thấy các sản phẩm gần như được mua cùng số lượng như nhau.

1. **So sánh phương thức thanh toán nào được sử dụng nhiều hơn giữa Ewallet và Credit card ?**

A screenshot of a graph

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá**: Dựa vào biểu đồ ta thấy được phương thức thanh toán Ewallet là phương thức được sử dụng nhiều hơn Credit Card, nhưng mức chênh lệch này là không lớn.

1. **Có sự chênh lệch nào đáng kể giữa mua sắm của nam và nữ không?**

A graph with red and blue rectangles

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá:**  Nhìn Biểu đồ ta thấy sự chênh lệch mua saswsm giữa nam và nữ không cs sự chênh lệnh cao, gần như là bằng nhau

1. **So sánh doanh thu giữa 3 tháng đầu năm 2019 và 3 tháng đầu năm 2020?**

A blue rectangular object with red numbers

Description automatically generated

| Effectiveness | Đạt | Không đạt |
| --- | --- | --- |
| Accuracy | x |  |
| Discriminability | x |  |
| Separability | x |  |
| Visual popout |  | x |
| Perceptual grouping | x |  |

**Đánh Giá:** Nhìn biều đồ ta thấy sự chênh lệch doanh thu giữa 3 tháng đầu năm 2019 và 2020 là rất lớn và theo số liệu, lượng mua hàng 3 tháng đầu năm 2020 rất thấp(14 khách hàng trên tổng 1017 khách hàng) nên doanh thu chênh lệch rất nhiều

# **Reference**

<https://drive.google.com/file/d/1TBRaHCOakafAyb9We9NvbkT-Dg-MPC9h/view>

<https://courses.ctda.hcmus.edu.vn/mod/folder/view.php?id=75246>

Link Drive Nhóm: <https://drive.google.com/drive/folders/1O4dJf9uT6xnCcpXAvv--1j97Kd1wpHXX?usp=sharing>