x

**ĐẠI HỌC QUỐC GIA**

**THÀNH PHỐ HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO GIAI ĐOẠN 1**

|  |  |  |
| --- | --- | --- |
| **GIẢNG VIÊN HƯỚNG DẪN** |  | **Tiết Gia Hồng** |
|  | **Nguyễn Trần Minh Thư** |
|  | **Phạm Minh Tú** |
| **LỚP HỌC PHẦN** | **:** | **CSC10108\_20HTTT** |

# Thông tin nhóm:

|  |  |  |  |
| --- | --- | --- | --- |
| ***Mã nhóm*** | ***MSSV*** | ***Họ và tên*** | ***Ghi chú*** |
| [20HTTT]  [NHÓM 8] | 19127348 | Bùi Công Danh |  |
| 20127561 | Nguyễn Hoài Mẫn |  |
| 20127503 | Dương Hiển Lê Hoàng |  |

# II. Phân tích nhu cầu các bên liên quan:

1. Ban quản lý

Ban quản lý thực hiện đưa ra các chiến lược tăng ngạch bán hàng bằng cách: − Các sản phẩm và số lượng mặt hàng bán chạy, bán ít ở từng quốc gia. − Giá cả của các loại mặt hàng. − Xu hướng mặt hàng bán chạy/bán ít qua từng thời điểm.

2. Nhân viên kho

Nhân viên kho thực hiện quản lý tồn kho thông qua:

− Các sản phẩm, loại sản phẩm và số lượng hiện tại có trong kho. − Xem xét các loại sản phẩm bán chạy và các loại sản phẩm bán chậm để thực hiện các chiến lược xuất nhập hàng phù hợp, tránh tình trạng kho không có các mặt hàng tiềm năng hoặc các sản phẩm bị ứ động.

1. Nhân viên chăm sóc khách hàng

Thông qua lịch sử mua hàng của các khách hàng, nhân viên chăm sóc có thể thực hiện chiến lược giữ chân người dùng thông qua:

− Các sản phẩm được khách hàng mua gần đây. − Sản phẩm được khách hàng mua nhiều nhất. − Sản phẩm có giá trị cao mà khách hàng đã mua.

Đề xuất nhu cầu các bên liên quan:

1. Ban quản lý

Ban quản lý muốn xem các nhân viên đã bán được bao nhiêu mặt hàng, doanh thu mà nhân viên nhận được từ việc bán hàng là bao nhiêu. Họ có thể dựa vào KPI để khen thưởng các cá nhân hoặc tập thể đã hoàn thành hoặc vượt trên mục tiêu đã đề ra.

1. Nhân viên kho

Nhân viên kho muốn xem thêm các loại mặt hàng nào bán chạy, bán thấp theo các thời điểm trong năm để thực hiện việc nhập hàng, chẳng hạn như phục vụ cho các sự kiện lớn (Black Friday, 11.11, …)

3. Nhân viên chăm sóc khách hàng

Xem lượt mua hàng để tạo ra các chương trình khuyến mãi cho khách hàng (voucher cho khách mới, thành viên vip mua hàng có chiết khấu) Ngoài ra, nhân viên cũng dựa trên số tiền khách hàng đã dành cho việc mua sấm để phân các cấp bậc khách hàng (bạc, vàng, …) để có các chiến lược chăm sóc khách hàng đối với từng loại cấp bậc tương ứng.

# III. Data profiling:

Tập dữ liệu gồm 8 cột, 541909 dòng

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | Description |
| Field Data type | String |
| Field length | 100 |
| Data profiling sumary statistic |  |
| NULL | 1454 |
| Missing | 0 |
| Actual | 540455 |
| Completeness | 99.73% |
| Cardinality | 4192 |
| Uniqueness | 0.77% |
| Distinctness | 0.78% |
| Data profiling additional statistic |  |
| Field length (MIN) | 1 |
| Field length (MAX) | 35 |
| Field value (MIN) | ? |
| Field value (MAX) | DECORATION WHITE CHICK MAGIC GARDEN |

Top 5 field values

|  |  |  |
| --- | --- | --- |
| Value | Count | Percentage |
| WHITE HANGING HEART T-LIGHT HOLDER | 2369 | 0.44% |
| REGENCY CAKESTAND 3 TIER | 2200 | 0.41% |
| JUMBO BAG RED RETROSPOT | 2159 | 0.40% |
| PARTY BUNTING | 1727 | 0.32% |
| LUNCH BAG RED RETROSPOT | 1638 | 0.30% |

Field format

|  |  |  |
| --- | --- | --- |
| No. | Pattern | Pattern Percentage |
| 1 | \p{L}+ \p{L}+ \p{L}+ \p{L}+ | 35% |
| 2 | (\p{L}+ )+ | 38% |
| 3 | .\* | 27% |

Top 5 field format value

|  |  |  |
| --- | --- | --- |
| Length | Count | Percentage |
| 30 | 41525 | 7.66% |
| 28 | 37100 | 6.85% |
| 31 | 35424 | 6.54% |
| 29 | 35226 | 6.50% |
| 27 | 35169 | 6.49% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | Country |
| Field Data type | String |
| Field length | 100 |
| Data profiling sumary statistic |  |
| NULL | 0 |
| Missing | 0 |
| Actual | 541909 |
| Completeness | 100% |
| Cardinality | 38 |
| Uniqueness | 0.007% |
| Distinctness | 0.007% |
| Data profiling additional statistic |  |
| Field length (MIN) | 3 |
| Field length (MAX) | 20 |
| Field value (MIN) | RSA |
| Field value (MAX) | United Arab Emirates |

Top 5 value

|  |  |  |
| --- | --- | --- |
| Country | Count | Percentage |
| United Kingdom | 495478 | 91.43% |
| Germany | 9495 | 1.75% |
| France | 8557 | 1.58% |
| EIRE | 8196 | 1.51% |
| Spain | 2533 | 0.47% |

Top 5 field format value

|  |  |  |
| --- | --- | --- |
| Length | Count | Percentage |
| 14 | 495508 | 91.44% |
| 7 | 13295 | 2.45% |
| 6 | 11694 | 2.16% |
| 4 | 8196 | 1.51% |
| 11 | 4819 | 0.89% |

Field format

|  |  |  |
| --- | --- | --- |
| No. | Pattern | Pattern Percentage |
| 1 | UNITED KINGDOM | 92% |
| 2 | \p{L}+ | 8% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | CustomerID |
| Field Data type | String |
| Field length | 5 |
| Data profiling sumary statistic |  |
| NULL | 135080 |
| Missing | 0 |
| Actual | 406829 |
| Completeness | 75.07% |
| Cardinality | 4372 |
| Uniqueness | 0.81% |
| Distinctness | 1.07% |
| Data profiling additional statistic |  |
| Field length (MIN) | 5 |
| Field length (MAX) | 5 |
| Field value (MIN) | 12346 |
| Field value (MAX) | 18287 |

Top 5 field value

|  |  |  |
| --- | --- | --- |
| ca | Count | Percentage |
| 17841 | 7983 | 1.47% |
| 14911 | 5903 | 1.09% |
| 14096 | 5128 | 0.95% |
| 12748 | 4642 | 0.86% |
| 14606 | 2782 | 0.51% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | Quanity |
| Field Data type | Int |
| Field length | 10 |
| Data profiling sumary statistic |  |
| NULL | 0 |
| Missing | 0 |
| Actual | 541909 |
| Completeness | 100% |
| Cardinality | 722 |
| Uniqueness | 0.13% |
| Distinctness | 0.13% |
| Data profiling additional statistic |  |
| Field length (MIN) | 2 |
| Field length (MAX) | 3 |
| Field value (MIN) | -12 |
| Field value (MAX) | 144 |

Top 5 field value:

|  |  |  |
| --- | --- | --- |
| Value | Count | Percentage |
| 1 | 148227 | 27.35% |
| 2 | 81829 | 15.10% |
| 12 | 61063 | 11.27% |
| 6 | 40868 | 7.54% |
| 4 | 38484 | 7.10% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | InvoiceDate |
| Field Data type | Date |
| Field length | 20 |
| Data profiling sumary statistic |  |
| NULL | 0 |
| Missing | 0 |
| Actual | 541909 |
| Completeness | 100% |
| Cardinality | 305 |
| Uniqueness | 0.06% |
| Distinctness | 0.06% |
| Data profiling additional statistic |  |
| Field data type | Date |
| Field length (MIN) | 9 |
| Field length (MAX) | 9 |
| Field value (MIN) | 12/1/2010 |
| Field value (MAX) | 12/9/2011 |
| Field format | Mm/dd/yyyy |

Top 5 field value:

|  |  |  |
| --- | --- | --- |
| Value | Count | Percentage |
| 12/5/2011 12:00:00 AM | 5331 | 0.98% |
| 12/8/2011 12:00:00 AM | 4940 | 0.91% |
| 11/29/2011 12:00:00 AM | 4313 | 0.80% |
| 11/16/2011 12:00:00 AM | 4195 | 0.77% |
| 11/11/2011 12:00:00 AM | 4089 | 0.75% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | InvoiceNo |
| Field Data type | String |
| Field length | 6 |
| Data profiling sumary statistic |  |
| NULL | 9291 |
| Missing | 0 |
| Actual | 532618 |
| Completeness | 98.29% |
| Cardinality | 22061 |
| Uniqueness | 4.17% |
| Distinctness | 4324% |
| Data profiling additional statistic |  |
| Field length (MIN) | 6 |
| Field length (MAX) | 6 |
| Field value (MIN) | 536592 |
| Field value (MAX) | 581492 |

Top 5 field value:

|  |  |  |
| --- | --- | --- |
| Value | Count | Percentage |
| 573585 | 1114 | 0.21% |
| 581219 | 749 | 0.14% |
| 581492 | 731 | 0.13% |
| 580729 | 721 | 0.13% |
| 558475 | 705 | 0.13% |

|  |  |
| --- | --- |
| Input metadata |  |
| Fieldname | StockCode |
| Field Data type | String |
| Field length | 6 |
| Data profiling sumary statistic |  |
| NULL | 0 |
| Missing | 0 |
| Actual | 541909 |
| Completeness | 100% |
| Cardinality | 3958 |
| Uniqueness | 0.73% |
| Distinctness | 0.73% |
| Data profiling additional statistic |  |
| Field length (MIN) | 1 |
| Field length (MAX) | 12 |
| Field value (MIN) | POST |
| Field value (MAX) | 15056N |

Top 5 field value:

|  |  |  |
| --- | --- | --- |
| Value | Count | Percentage |
| 85123A | 2380 | 0.44% |
| 22423 | 2203 | 0.41% |
| 85099B | 2159 | 0.40% |
| 47566 | 1727 | 0.32% |
| 20725 | 1639 | 0.30% |

Top 5 field format value:

|  |  |  |
| --- | --- | --- |
| Length | Count | Percentage |
| 5 | 487036 | 89.87% |
| 6 | 51488 | 9.50% |
| 4 | 1276 | 0.24% |
| 1 | 715 | 0.13% |
| 3 | 710 | 0.13% |
| 7 | 390 | 0.07% |

Format value:

|  |  |  |
| --- | --- | --- |
| No. | Pattern | Pattern Percentage |
| 1 | \d\d\d\d\d | 91% |
| 2 | \d\d\d\d\d\p{L} | 9% |

# IV. Task abstraction, Idiom và biểu đồ:

Câu 1: Top 10 mặt hàng có số lượng các mặt hàng bán chạy ở tất cả các quốc gia

- Lấy danh sách các mặt hàng bán của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng tất cả các mặt hàng

=>Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng từng loại mặt hàng giao dịch

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các sản phẩm có số lượng giao dịch nhiều nhất => 10 sản phẩm bán chạy nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 10 sản phẩm có số lượng giao dịch nhiều nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm được bán ra  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 10 |

Chart, bar chart

Description automatically generated

Câu 2: Top 10 mặt hàng có số khách hàng hủy đơn nhiều nhất

- Lấy danh sách các mặt hàng bị hủy đơn

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số sản phẩm giao dich bị hủy

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng mỗi loại sản phẩm giao dich bị hủy

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các tổng số sản phẩm có số lượng hủy giao dịch nhiều nhất => 10 mặt hàng có số khách hàng hủy đơn nhiều nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 10 sản phẩm có số lượng hủy giao dịch nhiều nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm giao dịch  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 10 |

Chart

Description automatically generated

Câu 3: Top 10 sản phẩm có doanh thu nhiều nhất trong 3 tháng cuối năm

- Lấy danh sách các mặt hàng bán của cửa hàng trong 3 tháng cuối năm

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số doanh thu của mỗi loại mặt hàng giao dịch trong 3 tháng cuối năm

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các sản phẩm có số doanh thu 3 tháng cuối năm => 10 sản phẩm doanh thu cao nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (doanh thu) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 10 sản phẩm có doanh thu cao nhất trong 3 tháng gần nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện doanh thu sản phẩm  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 10 |

Chart, bar chart

Description automatically generated

Câu 4: Top 10 sản phẩm có lượng bán ra nhiều nhất trong 3 tháng cuối năm

- Lấy danh sách các mặt hàng bán của cửa hàng trong 3 tháng cuối năm

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng của mỗi loại mặt hàng giao dịch trong 3 tháng cuối năm

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các sản phẩm có số giao dịch 3 tháng cuối năm => 10 sản phẩm có giao dịch cao nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 10 sản phẩm tổng số giao dịch nhiều nhất trong 3 tháng gần nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm giao dịch  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 10 |

Chart, bar chart

Description automatically generated

Câu 5: Top 10 sản phẩm có lượng bán ra ít nhất trong 3 tháng cuối năm

- Lấy danh sách các mặt hàng bán của cửa hàng trong 3 tháng cuối năm

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng của mỗi loại mặt hàng giao dịch trong 3 tháng cuối năm

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các sản phẩm có số giao dịch 3 tháng cuối năm => 10 sản phẩm có giao dịch cao nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 10 sản phẩm tổng số giao dịch ít nhất trong 3 tháng gần nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm giao dịch  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 10 |

Chart, bar chart

Description automatically generated

Câu 6: Top 10 khách hàng đã chi tiêu nhiều nhất cho cửa hàng

- Lấy danh sách các khách hàng của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số doanh thu của cửa hàng theo mỗi khách hàng

=>Action: annalyze=>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các khách hàng có số chi tiêu nhiều nhất => 10 khách hàng có chi tiêu cao nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (doanh thu) và 1 thuộc tính categorical (quốc gia) |
| Why: Task | Tìm 10 sản phẩm có doanh thu cao nhất trong 3 tháng gần nhất |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện doanh thu sản phẩm  + Cột Y thể hiện quốc gia  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo quốc gia  + Gắn nhãn doanh thu của cửa hàng cho từng quốc gia |
| Scale | Items: 10 |

Text

Description automatically generated

Câu 7: Tổng doanh thu của cửa hàng theo từng quốc gia

- Lấy danh sách các mặt hàng bán của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng từng loại sản phẩm bán theo từng quốc gia

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng doanh thu các sản phẩm theo từng quốc gia

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh số doanh thu từng quốc gia của cửa hàng => Số doanh thu nhiều nhất và ít của các quốc gia

=>search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (doanh thu) và 1 thuộc tính categorical (quốc gia) |
| Why: Task | Tìm doanh thu của các nước |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện doanh thu  + Cột Y thể hiện quốc gia  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 38 |

Chart

Description automatically generated with medium confidence

Câu 8: Top 20 sản phẩm có lượng bán ra nhiều nhất trong năm

- Lấy danh sách các mặt hàng bán của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số sản phẩm trong năm

=>Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng từng loại mặt hàng giao dịch trong năm

=>Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- So sánh các sản phẩm có số lượng giao dịch nhiều nhất => 20 sản phẩm bán chạy nhất

=>Action: search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 20 sản phẩm có số lượng giao dịch nhiều nhất trong năm |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm được bán ra  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 20 |

Chart, bar chart

Description automatically generated

Câu 9: Top 20 sản phẩm bị trả hàng lại nhiều nhất trong năm

Lấy danh sách các mặt hàng bán của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng từng loại mặt hàng giao dịch trong năm

=>Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbutio

- So sánh các sản phẩm có số lượng giao dịch trong năm => 20 sản phẩm bị trả hàng nhiều nhất

=>search=>browse=>query=>compare

|  |  |
| --- | --- |
| Idiom | Bar char |
| What: Data | Table: 1 giá trị quantitative (số lượng) và 1 thuộc tính categorical (sản phẩm) |
| Why: Task | Tìm 20 sản phẩm bị trả hàng lại nhiều nhất trong năm |
| How: Encode | - Mask: Area, line  - Channel:  + Cột X thể hiện số lượng sản phẩm được bán ra  + Cột Y thể hiện sản phẩm  + dán nhãn số lượng từng sản phẩm  + Sắp xếp giảm dần theo số lượng  + Màu: 1 màu xanh dương  + Căn chỉnh đều theo description  + Gắn nhãn số lượng cho từng sản phẩm |
| Scale | Items: 20 |

Application

Description automatically generated with low confidence

Câu.10 Top 5 tỷ lệ doanh thu quốc gia chi tiêu nhiều nhất cho cửa hàng

- Lấy danh sách các mặt hàng bán của cửa hàng

=>Action: search --> browse

=>Target: Attributes --> One -->Distrbution

- Tính tổng số lượng các sản phẩm bán theo từng quốc gia

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng doanh thu các sản phẩm theo từng quốc gia

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tổng doanh thu các sản phẩm của tất cả các quốc gia

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

- Tính tỷ lệ doanh thu của từng quốc gia của cửa hàng => Tỷ lệ 5 quốc gia chi tiêu nhiều nhất cho cửa hàng

=> Action: annalyze =>produce=>derive

=>Target: Attributes --> One -->Distrbution

|  |  |
| --- | --- |
| Idiom | Pie char |
| What: Data | Table: 1 giá trị quantitative (toal sales) và 1 thuộc tính categorical (quốc gia) |
| Why: Task | Tìm tỷ lệ top 5 doanh thu quốc gia có chi tiêu nhiều nhất |
| How: Encode | - Mask: Area  - Channel:  + Shape: Hình tròn  + Dán nhãn tỷ lệ phần tram doanh thu của quốc gia  + Màu: 6  + Căn chỉnh đều theo description |
| Scale | Items: 6 |

Chart, pie chart

Description automatically generated