Nguyễn Quốc Anh

**Pardus Consulting**

*Business Analytics Internship Test*

1. **Tổng quan**

Mục tiêu: Đánh giá chất lương, hiệu quả hoạt động của các trạm nhận giao

Dữ liệu: Bản Excel bao gồm **33** biến số (cột) và **22002** dòng dữ liệu (dòng), bao gồm

[1] "Trạm.lấy"

[2] "Trạm.giao"

[3] "Số.SO"

[4] "Số.PO"

[5] "Loại.hàng"

[6] "Trạng.thái"

[7] "TĐ.đặt.hàng"

[8] "TĐ.TMS.nhận.đơn"

[9] "Phương.thức.hẹn.giao.hàng"

[10] "OMS.Ngày.hẹn.giao.từ"

[11] "OMS.Ngày.hẹn.giao.đến"

[12] "Ngày.hẹn.giao.cuối.cùng.từ"

[13] "Ngày.hẹn.giao.cuối.cùng.đến"

[14] "Thời.điểm.cập.nhật.lấy.hàng.trạm.lấy"

[15] "Thời.điểm.nhập.kho"

[16] "Trạm.nhập.hàng.NCC"

[17] "Nhập.SC.Lấy"

[18] "Thời.điểm.nhập.kho.trạm.đích"

[19] "phân.công.giao.KH.lần.đầu"

[20] "phân.công.giao.KH.lần.cuối"

[21] "Thời.điểm.lưu.tạm.giao.hàng"

[22] "Lưu.hoàn.tất.cập.nhật.giao.hàng"

[23] "Nộp.tiền.thu.ngân"

[24] "Hình.thức.giao.hàng"

[25] "ĐV.giao.hàng"

[26] "Tỉnh.Thành.phố.giao.hàng"

[27] "Miền.giao.hàng"

[28] "Tỉnh.Thành.phố.lấy.hàng"

[29] "Miền.lấy.hàng"

[30] "Thành.tiền..VNĐ."

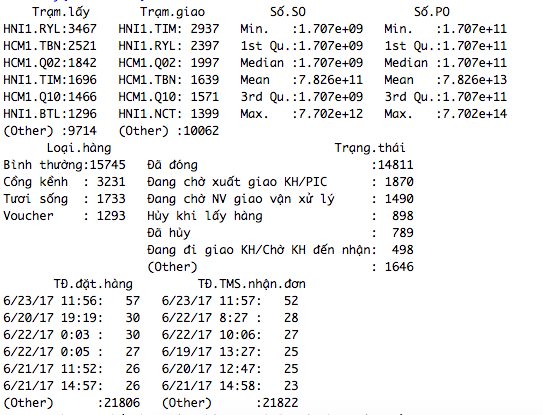
[31] "Luồng.SC.Nhập.trạm"

[32] "Mã.nguồn.đơn.hàng"

[33] "Nguồn.đơn.hàng"

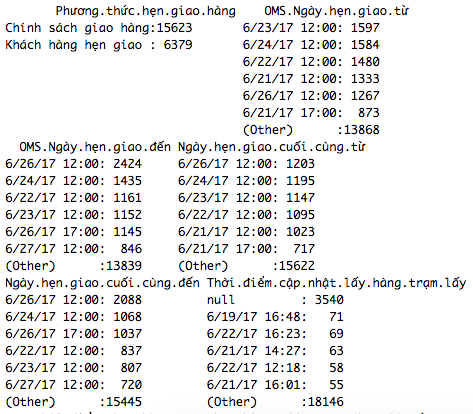
1. **Khám phá và miêu tả dữ liệu**

- Sau khi nhập dữ liệu vào R và phân tích cơ bản, đây là output của chương trình:

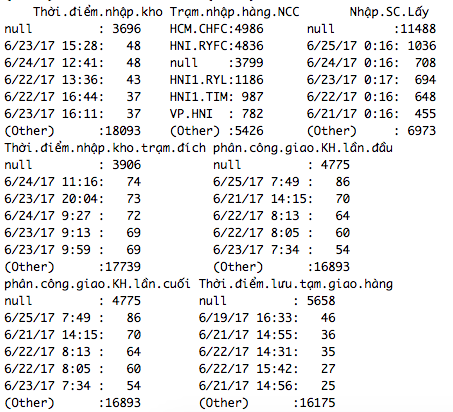


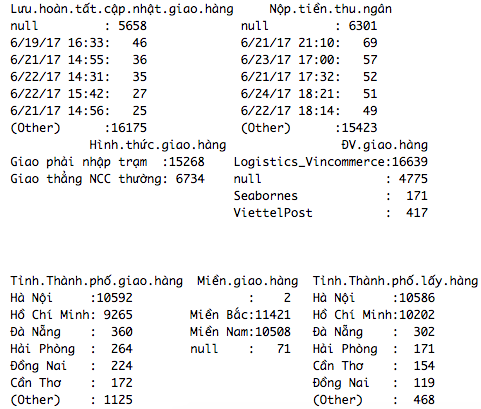
**Nhận xét**: Phần lớn những phân tích ở đây là tần số xuất hiện của các dữ liệu (như TĐ đặt hàng hay Trạng Thái). Ta nhận thấy Trạm lấy HNI1.RYL và trạm giao HNI1.TIM là xử lí nhiều đơn hàng nhất. Loại hàng Bình thường bao gồm 15745 đơn hàng, nhiều hơn tất cả thể loaị còn lại cộng vào (Cồng kềnh, Tươi Sống, và Voucher).

**\* Lưu ý**: Số PO và số SO dù là loại dữ liệu trừu tượng (Discrete), nhưng được lưu là Liên tiếp (Continuos) nên bị viết thành Scientic Connotation (1.0e+09,…). Điều nay không trở thành vấn đề, và có thể dễ dàng được giải quyết.

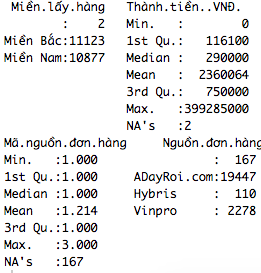


**Nhận xét**: 71% đơn hàng là theo Chính sách giao hàng, 29% còn lại là Khách hàng hẹn giao hàng. Phần lớn thông tin về thời điểm giao nhận hàng như trên.





**Nhận xét**: Giao phải nhập trạm chiếm 69.39% tất cả các hình thức giao hàng, với hình thức còn lại là Giao thẳng NCC thường. Hà Nội và HCM chiếm phần lớn cảc tỉnh thành phố giao nhận hàng. Về miền thì số đơn hàng được phân chia khá đồng đều.



**Nhận xét**: Phần lớn (88.39%) nguồn đơn hàng là từ AdayRoi.com. Tính cả Vinpro thì 98.74% đơn hàng được chuyển từ các công ty con của tập đoàn Vingroup. Đơn hàng lớn nhất trị giá 399,285,000 VNĐ, và trung bình là 2,360,064 VNĐ.

1. **Tính tỷ lệ số đơn hàng (SO) giao tại Hà Nội, Hồ Chí Minh và tỉnh khác**

|  |  |
| --- | --- |
| **2. Tính tỷ lệ số đơn hàng (SO) giao tại Hà Nội, Hồ Chí Minh và tỉnh khác** | |
| Hà Nội | 10592 |
| HCM | 9265 |
| Tỉnh khác | 2143 |

1. **Tính tỷ lệ số đơn hàng giao cùng miền và trái miền tại Hà Nội, Hồ Chí Minh và tỉnh khác**

|  |  |  |
| --- | --- | --- |
| **Tỉ lệ** | **Cùng Miền** | **Trái Miền** |
| **Hà Nội** | 42.51% | 5.64% |
| **HCM** | 38.08% | 4.03% |
| **Tỉnh khác** | 7.95% | 1.79% |

1. **Tính và visualize số đơn hàng có cùng trạm lấy, trạm giao và số đơn hàng khác trạm lấy, trạm giao của từng loại hàng**

|  |  |  |
| --- | --- | --- |
| **Loại hàng** | Cùng Trạm | Trái Trạm |
| **Bình thường** | 5540 | 10205 |
| **Cồng kềnh** | 3132 | 99 |
| **Tươi sống** | 1733 | 0 |
| **Voucher** | 0 | 1293 |

1. **Đánh giá chất lượng, hiệu quả hoạt động của các trạm lấy, giao**

**Tiêu chuẩn 1**: Ta so sánh thời gian nhận hàng và thời gian cam kết trước hàng sẽ được trả, sau đó dùng pivot table tìm xem bao nhiêu tỉ lệ bao lần mỗi trạm giao chuyển hàng muộn, từ đó có thể đánh giá xem trạm nào hoạt động hiệu quả.

|  |  |  |  |
| --- | --- | --- | --- |
| **Đơn hàng** | **Số lần muộn** | **Số đơn hàng** | **%Muộn** |
| HCM.CHFC | 375 | 636 | 58.96% |
| HCM1.GVP | 143 | 816 | 17.52% |
| HCM1.Q02 | 489 | 1997 | 24.49% |
| HCM1.Q07 | 430 | 1235 | 34.82% |
| HCM1.Q10 | 496 | 1571 | 31.57% |
| HCM1.TBN | 629 | 1639 | 38.38% |
| HCM1.TDC | 124 | 731 | 16.96% |
| HCM2.BHA | 46 | 118 | 38.98% |
| HCM2.BMT | 4 | 14 | 28.57% |
| HCM2.CTO | 32 | 92 | 34.78% |
| HCM2.ĐNG | 35 | 143 | 24.48% |
| HCM2.LXN | 9 | 28 | 32.14% |
| HCM2.NTG | 28 | 49 | 57.14% |
| HCS2.BLU | 0 | 4 | 0.00% |
| HCS2.KTM | 0 | 6 | 0.00% |
| HCS2.NTN | 0 | 1 | 0.00% |
| HNI.RYFC | 172 | 293 | 58.70% |
| HNI1.BTL | 292 | 1267 | 23.05% |
| HNI1.HDT | 167 | 930 | 17.96% |
| HNI1.LBN | 95 | 430 | 22.09% |
| HNI1.NCT | 281 | 1399 | 20.09% |
| HNI1.RAU | 130 | 130 | 100.00% |
| HNI1.RYL | 354 | 2397 | 14.77% |
| HNI1.TIM | 514 | 2937 | 17.50% |
| HNI2.HDG | 7 | 14 | 50.00% |
| HNI2.HLG | 2 | 14 | 14.29% |
| HNI2.HPG | 57 | 196 | 29.08% |
| HNI2.TBH | 6 | 26 | 23.08% |
| HNI2.VTI | 6 | 29 | 20.69% |
| HNS2.DPG | 0 | 2 | 0.00% |
| HNS2.NBH | 1 | 6 | 16.67% |
| VP.AGN | 1 | 4 | 25.00% |
| VP.CLH | 13 | 34 | 38.24% |
| VP.CTO | 38 | 55 | 69.09% |
| VP.DLK | 14 | 51 | 27.45% |
| VP.ĐNG | 133 | 222 | 59.91% |
| VP.DNI | 17 | 47 | 36.17% |
| VP.HCM | 784 | 1035 | 75.75% |
| VP.HGG | 0 | 2 | 0.00% |
| VP.HLG | 66 | 113 | 58.41% |
| VP.HNI | 389 | 984 | 39.53% |
| VP.HPG | 17 | 52 | 32.69% |
| VP.KGG | 1 | 3 | 33.33% |
| VP.TBH | 35 | 69 | 50.72% |
| VP.TQG | 12 | 14 | 85.71% |
| VP.TVH | 1 | 1 | 100.00% |
| VP.VLG | 4 | 4 | 100.00% |
| VP.VTI | 3 | 22 | 13.64% |
| VP.XK | 11 | 50 | 22.00% |
| VP.YBI | 64 | 71 | 90.14% |
| VP+ DNI | 11 | 19 | 57.89% |

**Tiêu chuẩn 2:** Ta coi mỗi đơn huỷ là lỗi trạm giao, ta có thể tìm tỉ lệ huỷ đơn hàng mỗi trạm

|  |  |
| --- | --- |
| **Trạm Giao** | **% Huỷ** |
| HCM.CHFC | 32.05% |
| HCM1.GVP | 7.26% |
| HCM1.Q02 | 21.99% |
| HCM1.Q07 | 8.47% |
| HCM1.Q10 | 3.89% |
| HCM1.TBN | 4.84% |
| HCM1.TDC | 6.11% |
| HCM1.VCR | 8.13% |
| HCM2.BHA | 13.46% |
| HCM2.BMT | 0.00% |
| HCM2.CTO | 2.04% |
| HCM2.ĐNG | 1.25% |
| HCM2.LXN | 0.00% |
| HCM2.NTG | 0.00% |
| HCS2.BLU | 0.00% |
| HCS2.KTM | 0.00% |
| HCS2.NTN | 0.00% |
| HNI.RYFC | 11.49% |
| HNI1.BTL | 3.78% |
| HNI1.HDT | 3.46% |
| HNI1.LBN | 4.60% |
| HNI1.NCT | 2.28% |
| HNI1.RAU | 0.77% |
| HNI1.RYL | 5.51% |
| HNI1.TIM | 3.42% |
| HNI1.VCR | 12.38% |
| HNI2.HLG | 33.33% |
| HNI2.HPG | 20.17% |
| HNI2.TBH | 0.00% |
| HNI2.VTI | 11.11% |
| HNS2.DPG | 0.00% |
| HNS2.NBH | 16.67% |
| VP.AGN | 0.00% |
| VP.CLH | 2.94% |
| VP.CTO | 0.00% |
| VP.DLK | 7.84% |
| VP.ĐNG | 4.50% |
| VP.DNI | 2.13% |
| VP.HCM | 6.96% |
| VP.HGG | 0.00% |
| VP.HLG | 0.00% |
| VP.HNI | 4.97% |
| VP.HPG | 1.92% |
| VP.KGG | 0.00% |
| VP.TBH | 0.00% |
| VP.TQG | 0.00% |
| VP.TVH | 0.00% |
| VP.VLG | 0.00% |
| VP.VTI | 0.00% |
| VP.XK | 6.00% |
| VP.YBI | 8.45% |
| VP+ DNI | 0.00% |

Note: Nếu em hiểu sai data, cho em xin lỗi. Em không hiểu 100% các viết tắt, tiêu đề các cột, và sự khác biệt giữa các mốc thời gian. Em cố phân tích theo ý hiểu cuả em.