

# **ROUTE AND TIME PREDICTION FOR DELIVERY OPTIMIZATION IN LOGISTICS**

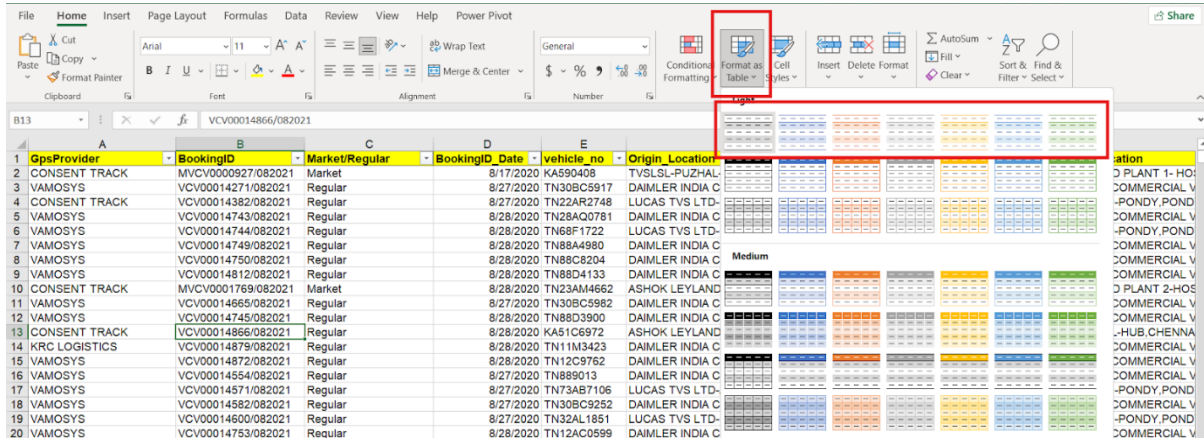
**--- Pham Dan Quyen ---**

## 1. Xác định các biến của dữ liệu

TÊN BIẾN	GIẢI THÍCH
GpsProvider	Đơn vị cung cấp GPS
BookingID	Mã đặt hàng
Market/Regular	Loại hợp đồng (Khách lẻ/Khách ký hợp đồng dài hạn)
BookingID_Date	Ngày đặt hàng
vehicle_no	Biển số xe
Origin_Location	Địa điểm bắt đầu
Destination_Location	Địa điểm đến
Org_lat_lon	Kinh độ của điểm xuất phát
Des_lat_lon	Kinh độ của điểm đến
Data_Ping_time	Dữ liệu thời gian trễ
Planned_ETA	Thời gian dự kiến
Current_Location	Địa điểm hiện tại
DestinationLocation	Địa điểm đến
actual_eta	Thời gian đến
Curr_lat	Vĩ độ hiện tại
Curr_lon	Kinh độ hiện tại

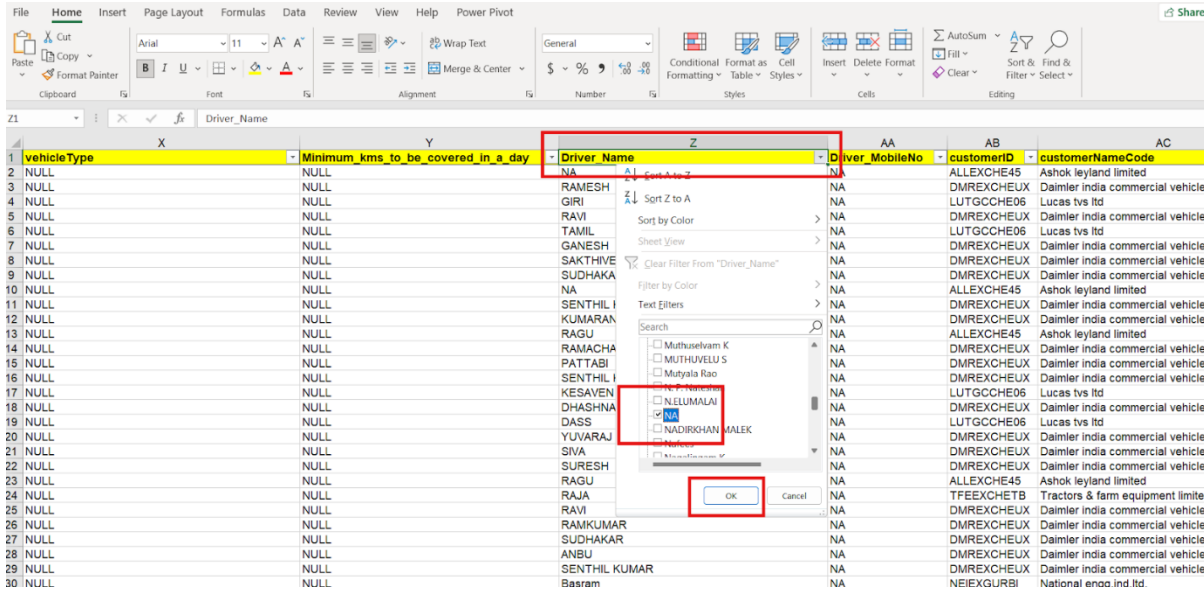
ontime	Ghi nhận đúng hẹn
delay	Ghi nhận trễ
OriginLocation_Code	Mã địa điểm bắt đầu
DestinationLocation_Code	Mã địa điểm đến
trip_start_date	Thời gian bắt đầu
trip_end_date	Thời gian đến nơi
TRANSPORTATION_DISTANCE_IN_KM	Số KM của hành trình
vehicleType	Kiểu phương tiện
Minimum_kms_to_be_covered_in_a_day	Số Km tối thiểu cần đi trong một ngày
Driver_Name	Tên tài xế
Driver_MobileNo	Số điện thoại tài xế
customerID	Tên khách hàng
customerNameCode	Mã khách hàng
supplierID	Tên nhà cung cấp
supplierNameCode	Mã nhà cung cấp
Material Shipped	Vật liệu vận chuyển

## 2. Xử lý dữ liệu trên Excel



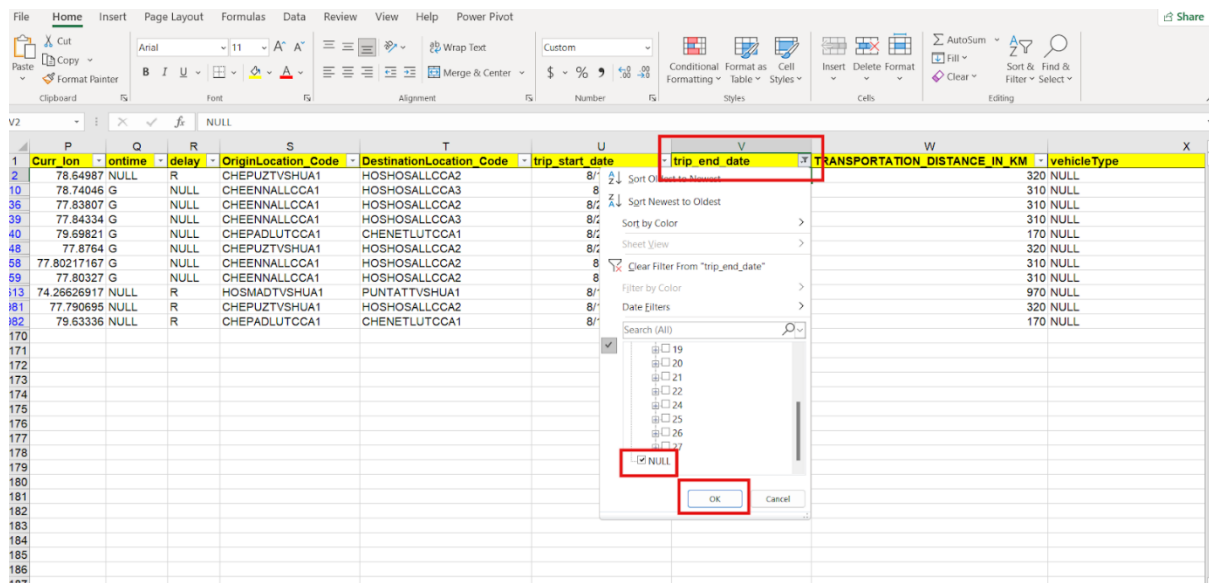
GpsProvider	BookingID	Market/Regular	BookingID	Date	vehicle_no	Origin_Location	Destination
CONSENT TRACK	MVCV0000927/082021	Market	8/17/2020	KA590408	TVSLSL-PUZHAL		3 PLANT 1-HO
VAMOSYS	VCV00014271/082021	Regular	8/27/2020	TN30BC5917	DAIMLER INDIA C		COMMERCIAL V
CONSENT TRACK	VCV00014382/082021	Regular	8/27/2020	TN22AR2748	LUCAS TVS LTD-		PONDY_POND
VAMOSYS	VCV00014743/082021	Regular	8/28/2020	TN28AQ0781	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014744/082021	Regular	8/28/2020	TN68F1722	LUCAS TVS LTD-		PONDY_POND
VAMOSYS	VCV00014749/082021	Regular	8/28/2020	TN88A4980	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014750/082021	Regular	8/28/2020	TN88C8204	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014812/082021	Regular	8/28/2020	TN88D4133	DAIMLER INDIA C		COMMERCIAL V
CONSENT TRACK	MVCV0001769/082021	Market	8/28/2020	TN23AM4662	ASHOK LEYLAND		3 PLANT 2-HOS
VAMOSYS	VCV00014665/082021	Regular	8/27/2020	TN30BC5982	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014745/082021	Regular	8/28/2020	TN88D3900	DAIMLER INDIA C		COMMERCIAL V
CONSENT TRACK	VCV00014866/082021	Regular	8/28/2020	KA51C6972	ASHOK LEYLAND		HUB, CHENNA
KRC LOGISTICS	VCV00014879/082021	Regular	8/28/2020	TN11M3423	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014872/082021	Regular	8/28/2020	TN12C9762	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014654/082021	Regular	8/27/2020	TN889013	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014571/082021	Regular	8/27/2020	TN73AB7106	LUCAS TVS LTD-		PONDY_POND
VAMOSYS	VCV00014582/082021	Regular	8/27/2020	TN30BC9252	DAIMLER INDIA C		COMMERCIAL V
VAMOSYS	VCV00014600/082021	Regular	8/27/2020	TN32AL1851	LUCAS TVS LTD-		PONDY_POND
VAMOSYS	VCV00014753/082021	Regular	8/28/2020	TN12AC0599	DAIMLER INDIA C		COMMERCIAL V

Hình 1: Tạo Table cho dữ liệu



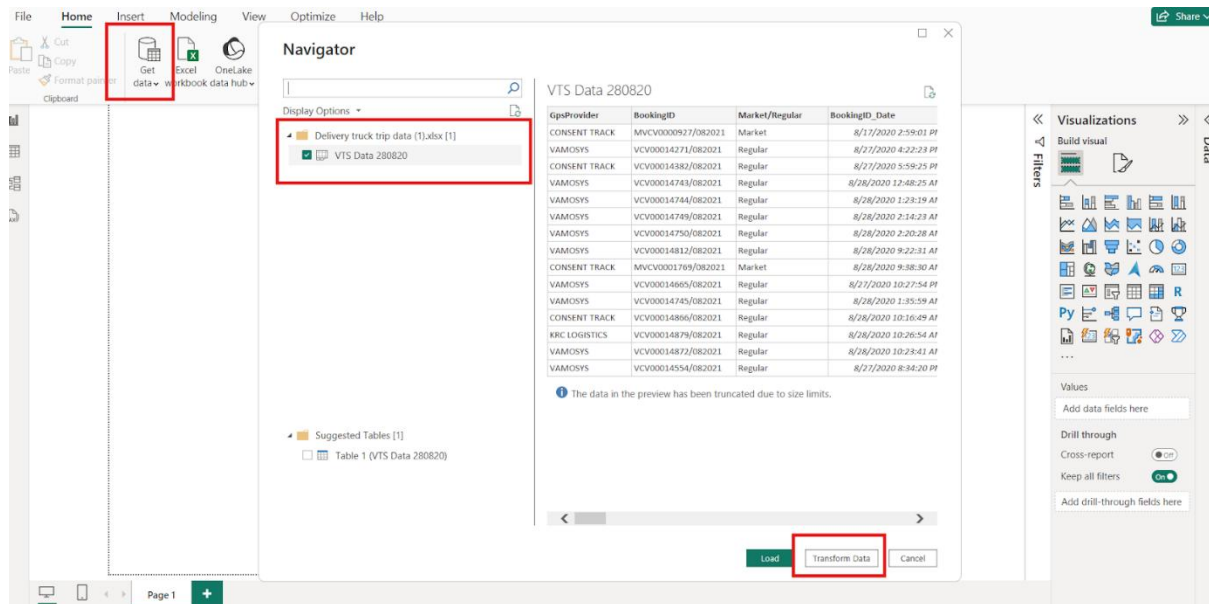
vehicleType	Minimum kms to be covered in a day	Driver Name	Driver MobileNo	customerID	customerNameCode
NULL	NULL	NA	ALLEXCHE45	Ashok leyland limited	
NULL	NULL	RAMESH	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	GIRI	LUTGCCHE06	Lucas tvs ltd	
NULL	NULL	RAVI	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	TAMIL	LUTGCCHE06	Lucas tvs ltd	
NULL	NULL	GANESH	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SAKTHIVE	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SUDHAKA	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	NA	ALLEXCHE45	Ashok leyland limited	
NULL	NULL	SENTHIL	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	KUMARAN	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	RAGU	ALLEXCHE45	Ashok leyland limited	
NULL	NULL	RAMACHA	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	PATTABI	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SENTHIL	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	KESAVEN	LUTGCCHE06	Lucas tvs ltd	
NULL	NULL	DHASHNA	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	DASS	LUTGCCHE06	Lucas tvs ltd	
NULL	NULL	YUVARAJ	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SIVA	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SURESH	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	RAGU	ALLEXCHE45	Ashok leyland limited	
NULL	NULL	RAJA	TFEEXCHETB	Tractors & farm equipment limited	
NULL	NULL	RAVI	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	RAMKUMAR	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SUDHAKAR	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	ANBU	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	SENTHIL KUMAR	DMREXCHEUX	Daimler india commercial vehicle	
NULL	NULL	Basram	NEIEXGURBI	National enoa ind ltd	

Hình 2: Xóa dòng dữ liệu không có thông tin Tên tài xế

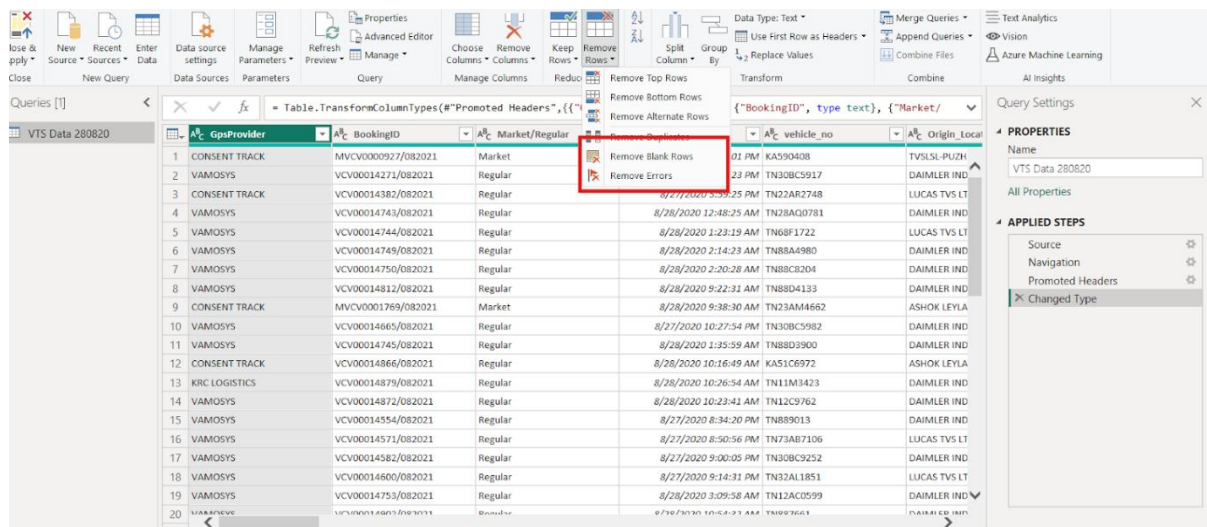


Hình 3: Xóa dòng dữ liệu không có thời gian kết thúc hành trình

### 3. Xử lý dữ liệu trên BI:

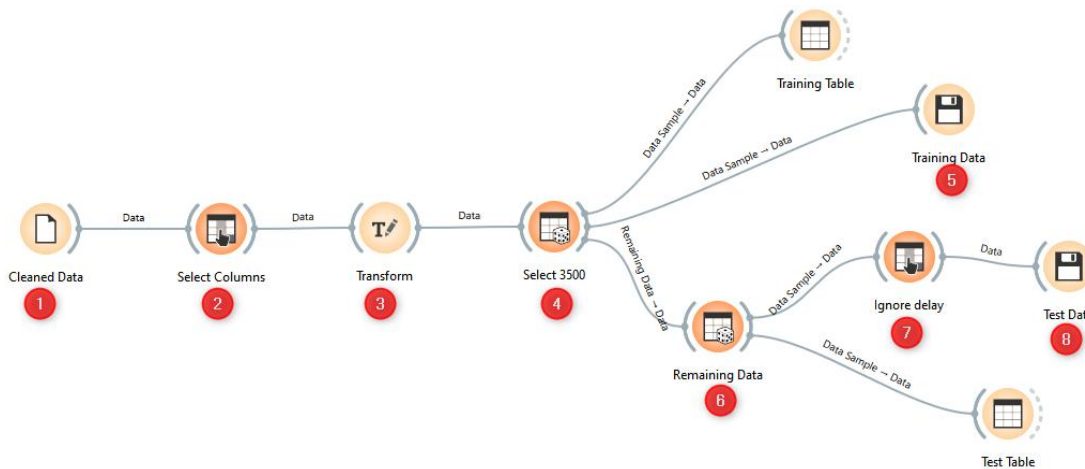


Hình 4: Chuyển dữ liệu từ Excel lên BI



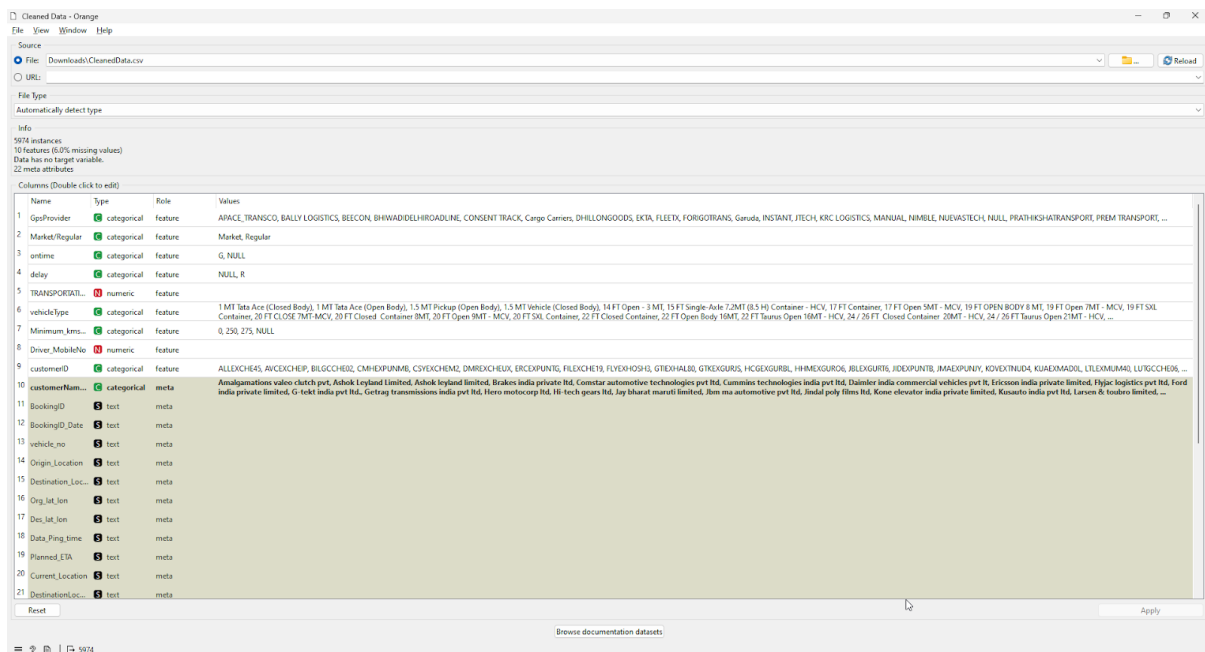
Hình 5: Xóa các cột trống trước khi đưa lên BI

## 4. Xử lý và chuẩn bị dữ liệu trên Orange



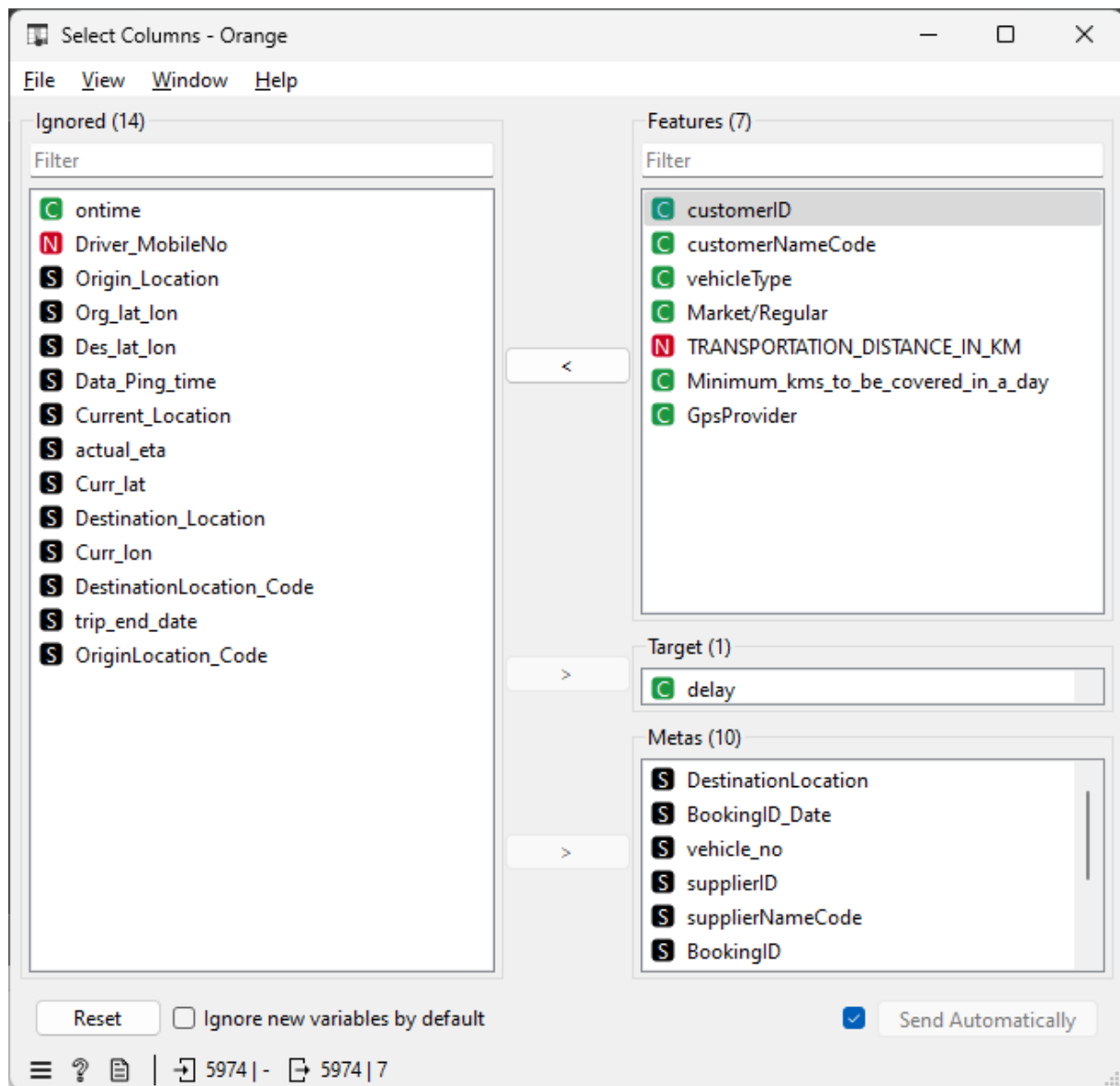
Hình 6: Mô hình phân tích dữ liệu trong Orange

Bước 1: Input data đã được xử lý ở excel.



Hình 7: Chỉnh sửa giá trị cho delay

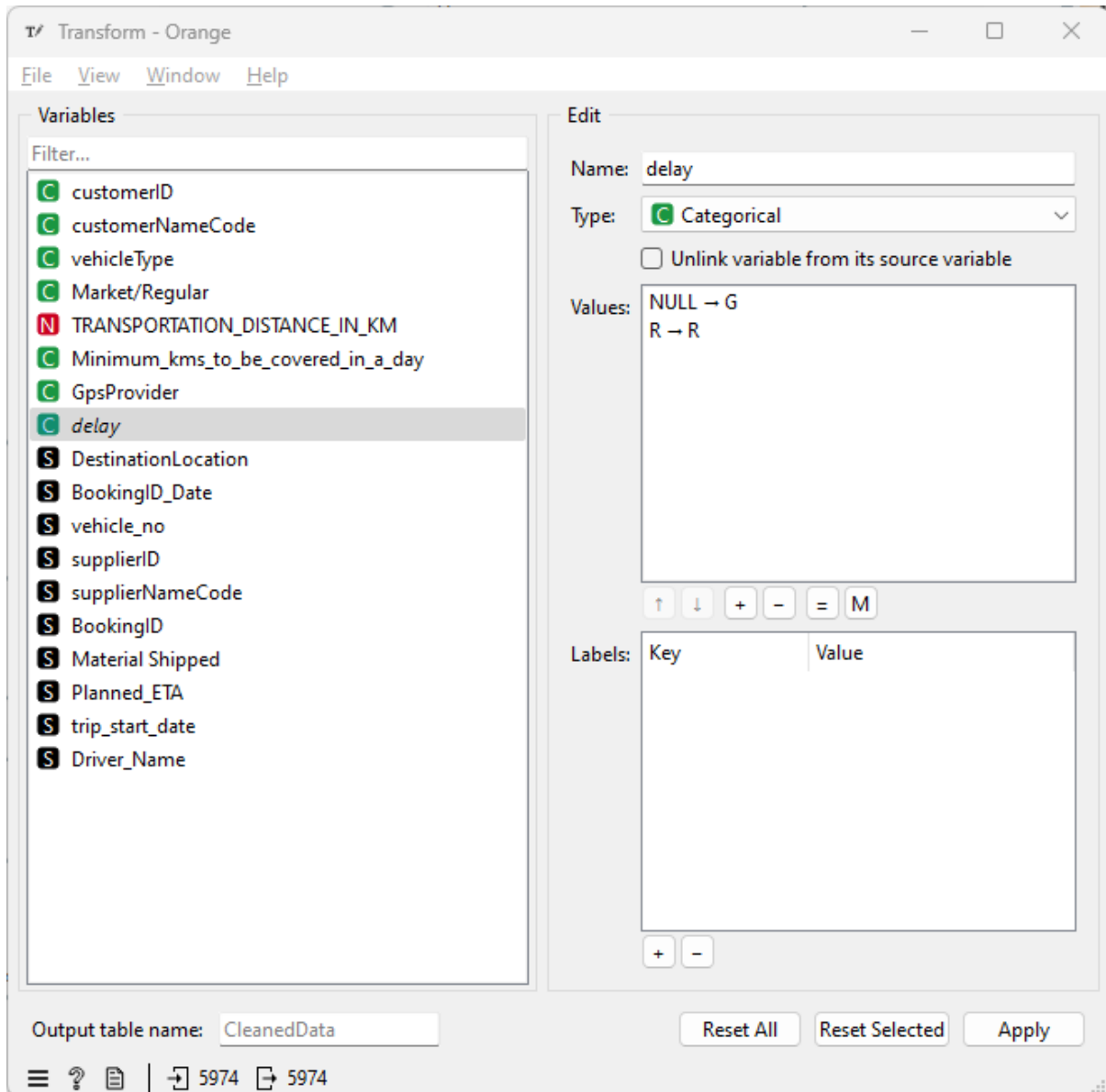
Bước 2: Chọn Delay làm biến mục tiêu. Ngoài ra còn lựa chọn 1 số Features, Metas và loại bỏ 1 số cột không phù hợp cho việc phân lớp.



Hình 8: Chọn Delay làm biến mục tiêu

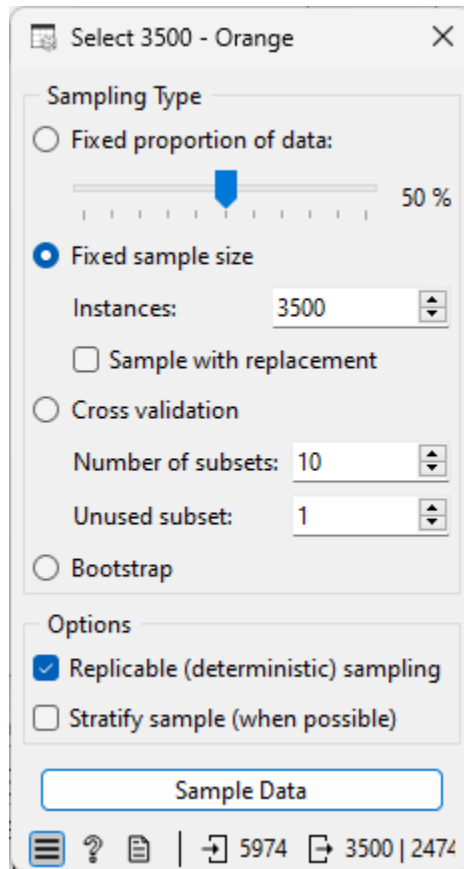


Bước 3: Chỉnh sửa giá trị cho delay. Null -> G (Green), R -> R (Red)



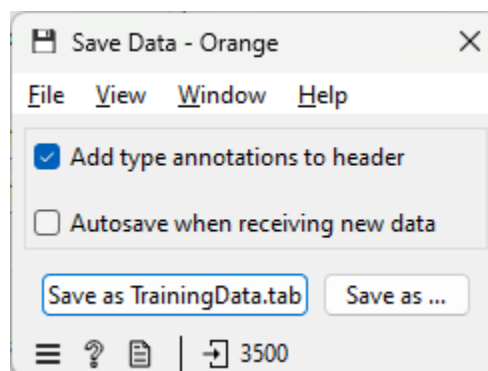
Hình 9: Chỉnh sửa giá trị cho delay

Tổng có 5974 dòng, Lấy 3500 data để sử dụng cho việc training phân lớp



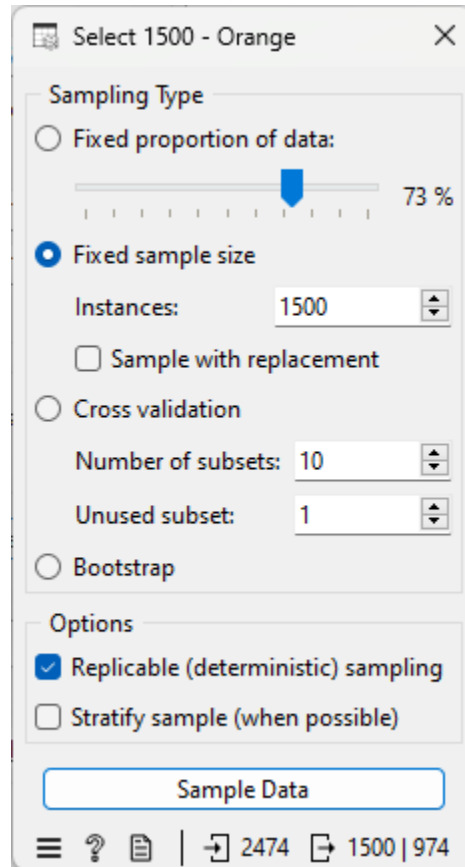
Hình 10: Kết quả chỉnh sửa giá trị Delay

Bước 4: Lưu Training data được lựa chọn dưới dạng .tab



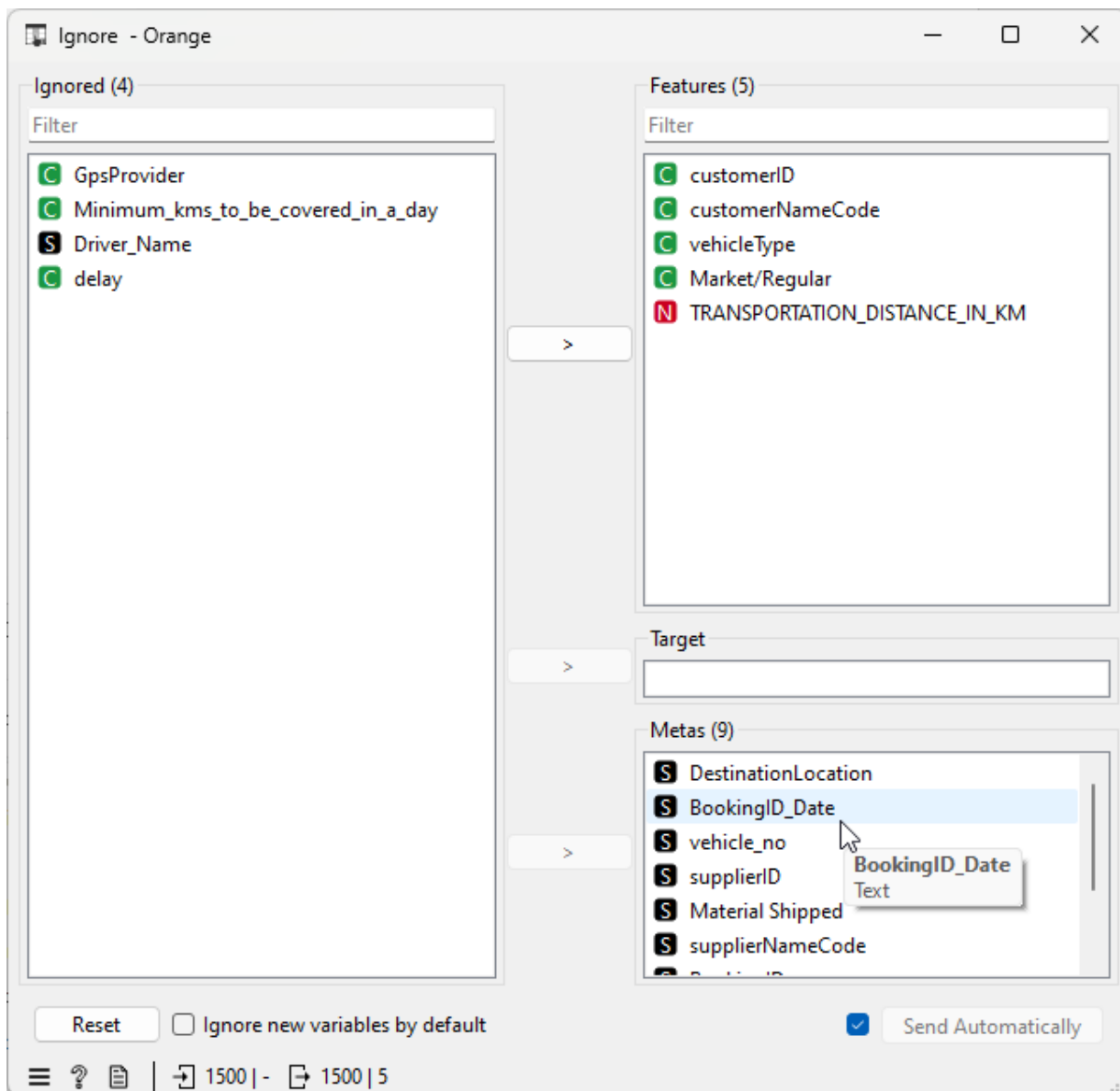
Hình 11: Lưu training data dưới dạng .tab

Bước 5: Lấy 1500 data từ phần còn lại để sử dụng cho việc dự đoán



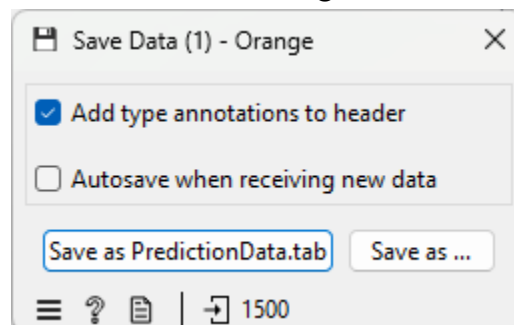
Hình 12: Lấy 1500 data từ phần còn lại để sử dụng cho việc dự đoán

Bước 6: Loại bỏ cột delay để dự đoán được chính xác.



Hình 13: Loại bỏ cột delay để dự đoán được chính xác

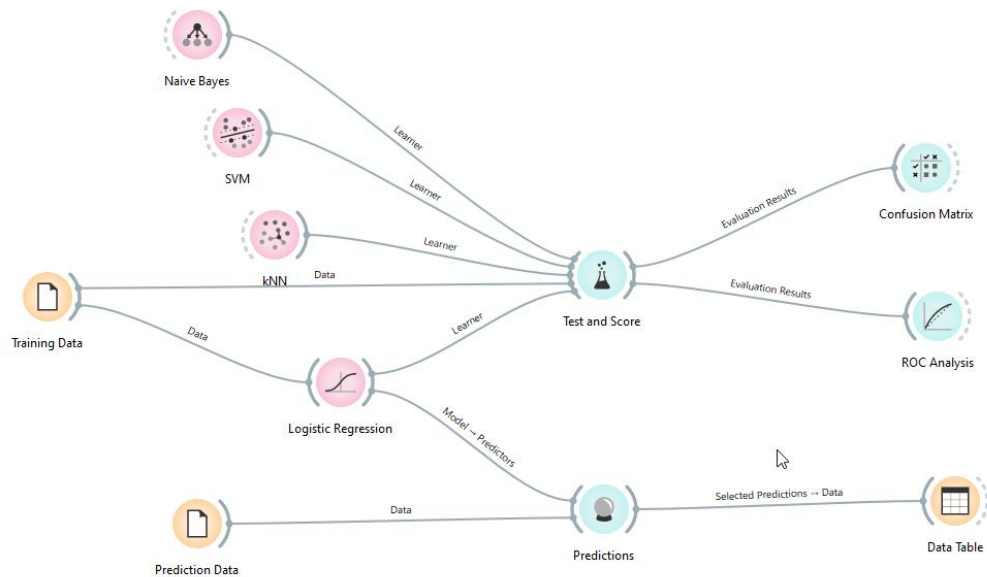
Bước 7: Lưu Training data được lựa chọn dưới dạng .tab



Hình 14: Lưu Training data được lựa chọn dưới dạng .tab

## 5. Hiển thị kết nối trong Orange

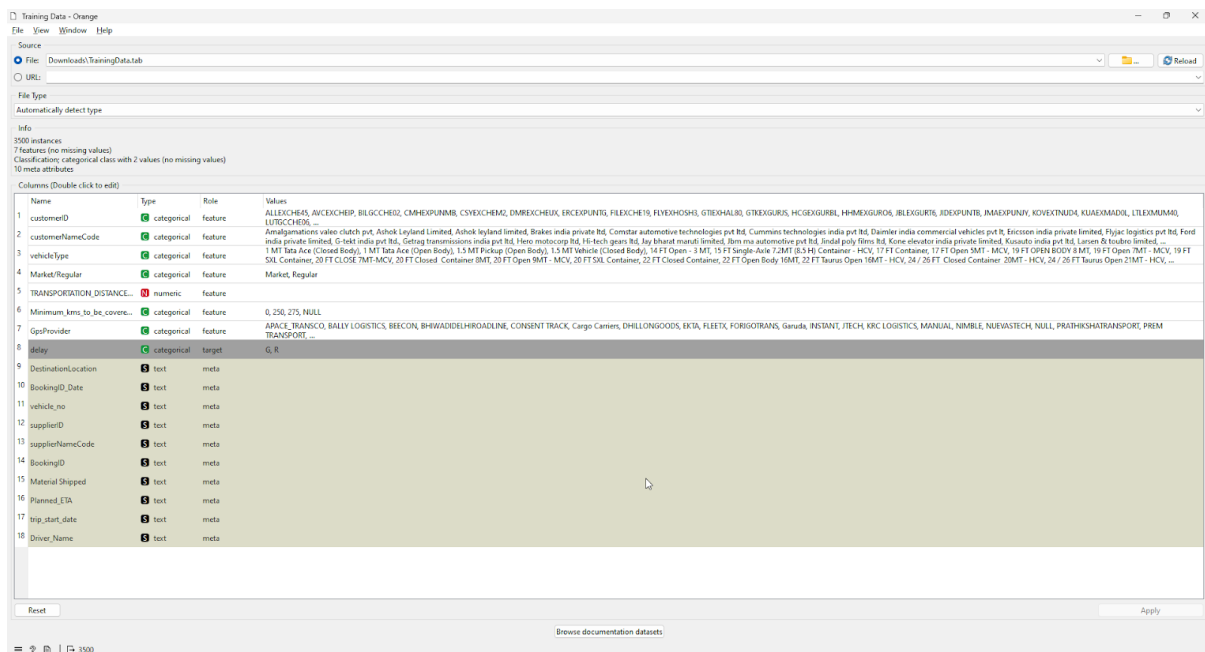
Sử dụng Orange cho quá trình phân lớp



Hình 15: Hiển thị kết nối trong Orange

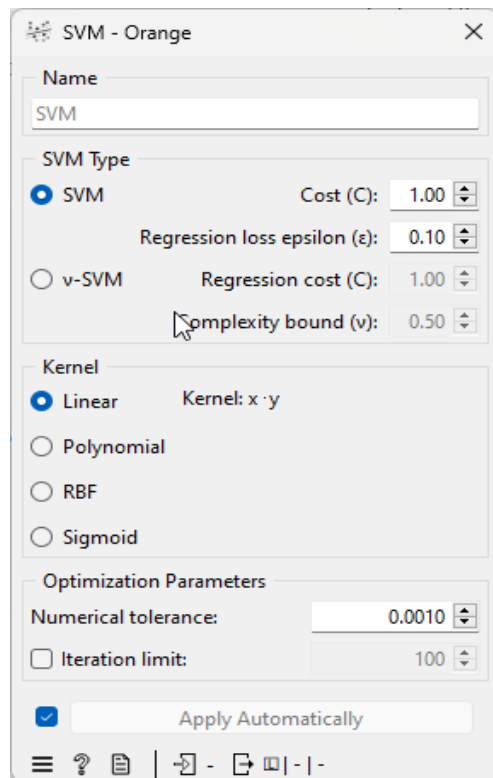
## 6. Các bước và model được thử nghiệm cho việc phân lớp

- Import data huấn luyện vào Orange. Set Delay làm target, các biến CustomerID, CustomerNameCode, VehicleType, Market/Regular, TRANSPORTATION\_DISTANCE\_IN\_KM, Minimum\_kms\_to\_be\_covered\_in\_a\_day, GpsProvider là biến chức năng và các biến còn lại siêu dữ liệu



Hình 16: Import data huấn luyện vào Orange

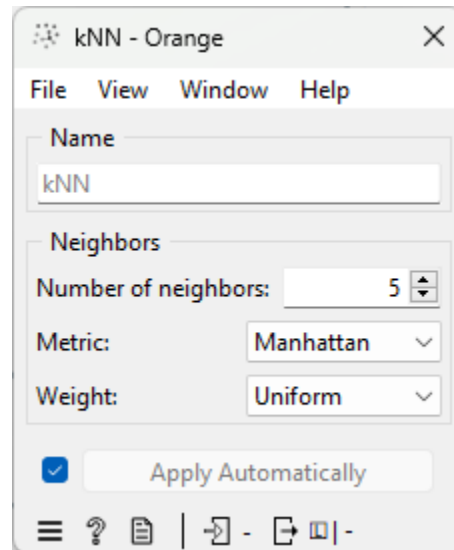
- **Naive Bayes:** Một mô hình phân loại xác suất nhanh và đơn giản dựa trên định lý Bayes với giả định rằng các đặc trưng là độc lập với nhau.



Hình 17: SVM

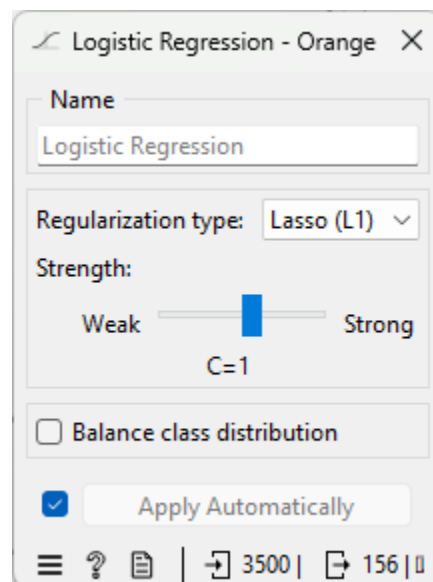
- **SVM:** Support Vector Machines ánh xạ đầu vào sang không gian đặc trưng có chiều cao hơn.

- kNN: Dự đoán dựa trên các điểm dữ liệu huấn luyện gần nhất. Widget kNN sử dụng thuật toán k-nearest neighbors để tìm kiếm k ví dụ huấn luyện gần nhất trong không gian đặc trưng và sử dụng giá trị trung bình của chúng làm dự đoán.



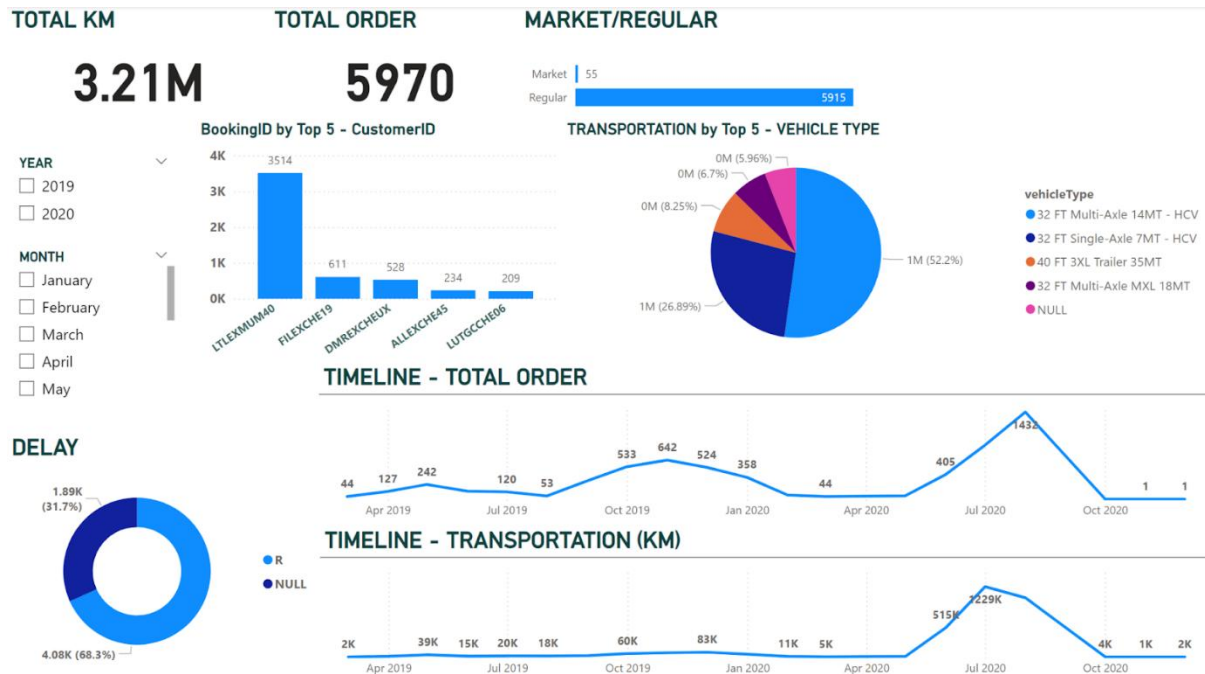
Hình 18: kNN

- Logistic Regression: Thuật toán phân loại hồi quy logistic với regularization LASSO (L1) hoặc ridge (L2). Hồi quy Logistic học một mô hình Logistic Regression từ dữ liệu và chỉ hoạt động cho các bài toán phân lớp.



Hình 19: Logistic Regression

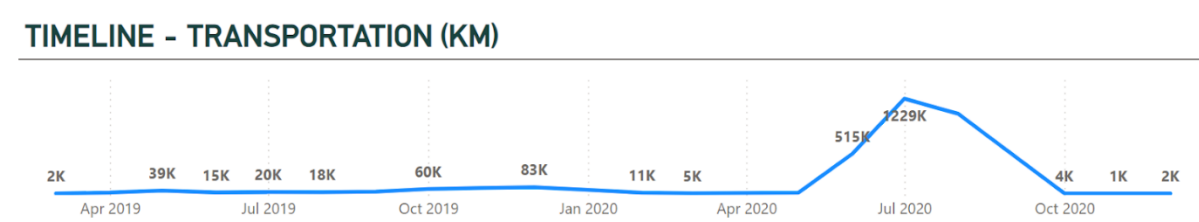
## 7. Tổng hợp kết quả từ Dashboard (Power BI)



Hình 20: Bảng kết quả tổng hợp của doanh nghiệp trong giai đoạn 2019 - 2020

- Tổng khối lượng vận chuyển:** Trong giai đoạn 2019-2020 tổng quãng đường vận chuyển đạt 3,21 triệu km phản ánh hiệu quả vận hành và khả năng đáp ứng nhu cầu thị trường một cách ổn định.

Đặc biệt giai đoạn từ tháng 04.2020 đến 07.2020 khối lượng lập đỉnh với mức 1229k km.

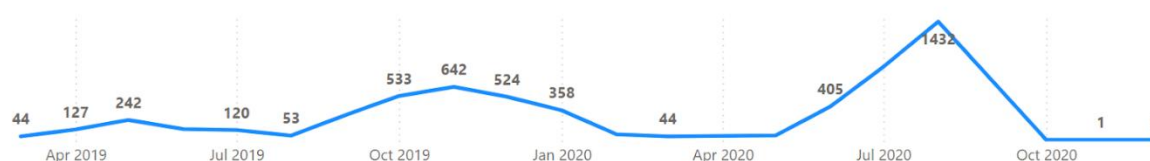


Hình 21: Kết quả khối lượng lập đỉnh từ tháng 4/2020 đến 07/2020

- Tổng số đơn hàng (Theo bookingID):** Trong giai đoạn 2019-2020 tổng đơn hàng đạt 5970 đơn hàng. Số lượng đơn hàng xử lý trong kỳ cho thấy năng lực quản lý đơn hàng của hệ thống đã đang đáp ứng được sức tăng trưởng của doanh nghiệp. Đặc biệt ở giai đoạn từ tháng 04.2020 đến 07.2020



## TIMELINE - TOTAL ORDER



Hình 22: Khối lượng đơn hàng giai đoạn 2019 - 2020

- **Số lượng đơn hàng Market/Regular:** Trong giai đoạn 2019-2020 tỷ lệ phân bổ giữa đơn hàng Market và Regular cho thấy có sự đa dạng trong tệp khách hàng, doanh nghiệp cần cân đối trong việc phục vụ giữa các nhóm khách hàng khác nhau.
- Năm 2019: 100% khách hàng là phân loại Regular

## MARKET/REGULAR

Regular

2720

Hình 23: Kết quả phân loại năm 2019

- Năm 2020: 98.22% khách hàng là phân loại Regular, 1.78% là phân loại Market

## MARKET/REGULAR

Market | 55

Regular

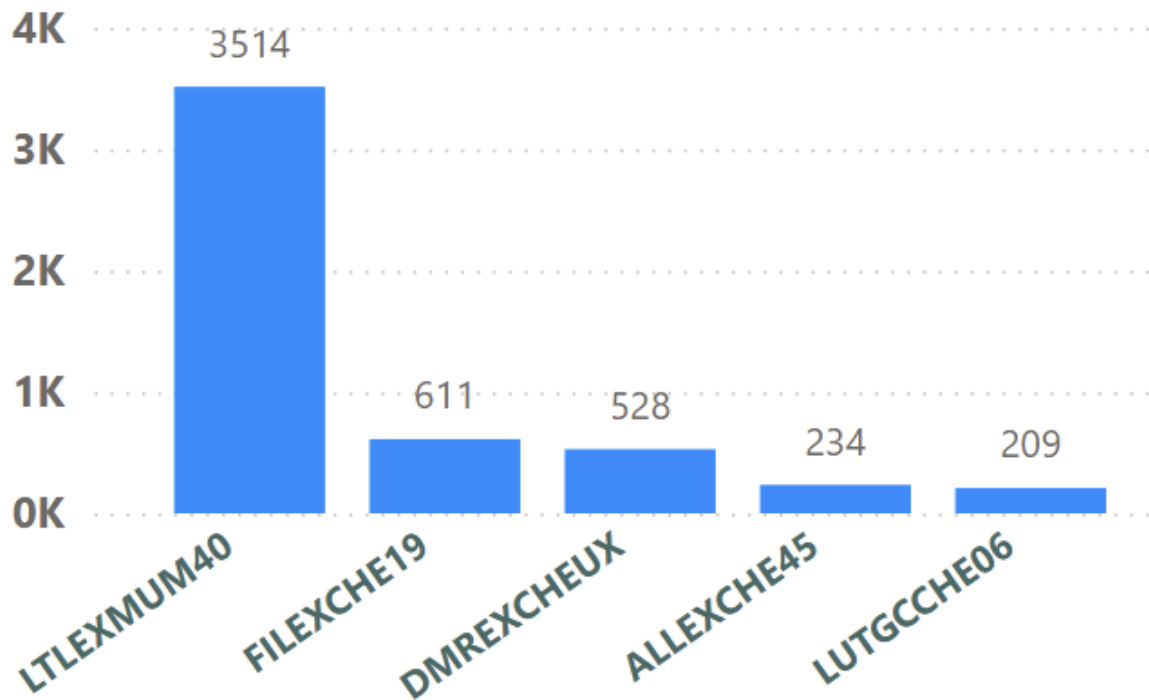
5915

Hình 24: Kết quả phân loại năm 2020

- **Top 5 khách hàng theo số lượng đơn hàng:** Trong giai đoạn 2019-2020 nhóm khách hàng chiến lược này chiếm tỷ trọng đáng kể trong tổng số đơn hàng, là đối

tượng cần tập trung duy trì và gia tăng sự hài lòng thông qua dịch vụ cá nhân hóa và ưu đãi đặc biệt.

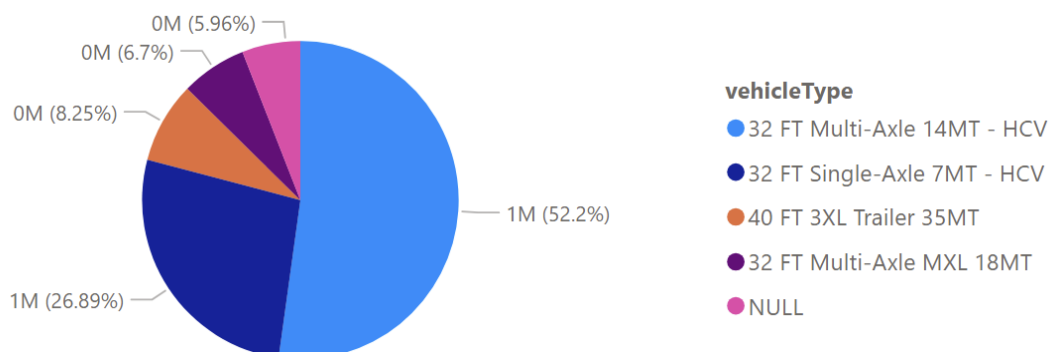
## BookingID by Top 5 - CustomerID



Hình 25: Top 5 loại hình vận chuyển theo quãng đường

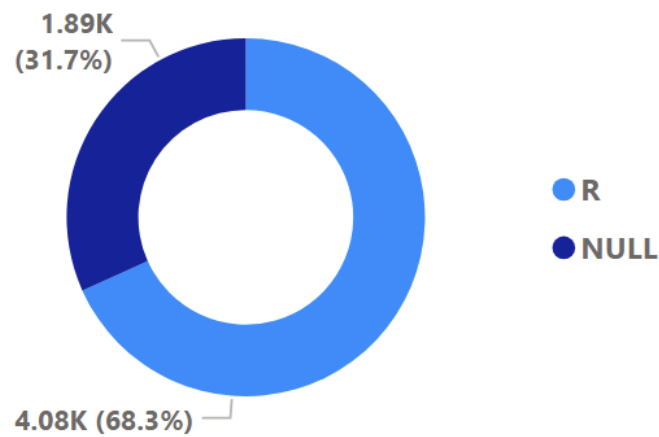
- Top 5 loại hình vận chuyển theo quãng đường vận chuyển:** Trong giai đoạn 2019-2020 kết quả cho thấy các loại hình vận chuyển được ưu tiên lựa chọn bởi khách hàng. Điều này nhấn mạnh vai trò quan trọng của việc quản lý phương tiện và lựa chọn loại hình phù hợp đối với từng nhóm nhu cầu khác nhau của khách hàng.

## TRANSPORTATION by Top 5 - VEHICLE TYPE



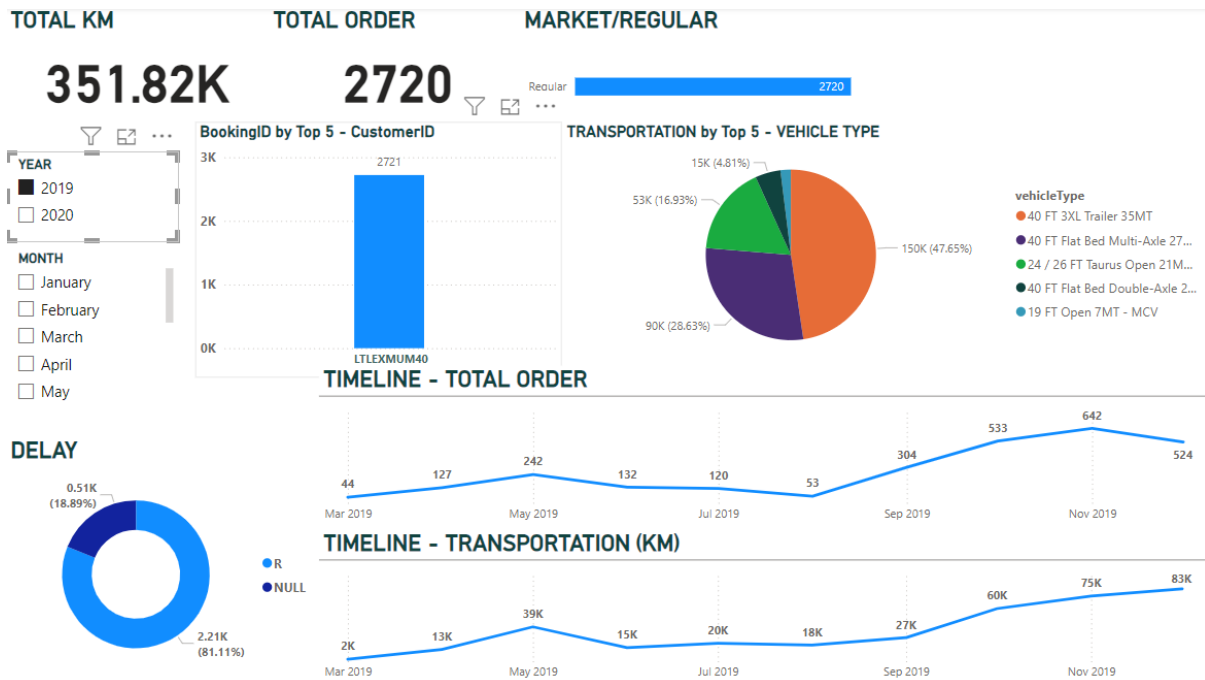
- **Tỷ lệ Delay:** Trong giai đoạn 2019-2020, tỷ lệ Delay (R) chiếm đến 68,3%, đây là một tỷ trọng khá cao đối với mô hình kinh doanh vận chuyển. Dù tỷ lệ chậm trễ nằm trong mức chấp nhận được, nhưng đây vẫn là lĩnh vực cần cải thiện để nâng cao trải nghiệm khách hàng và giảm thiểu khiếu nại.

## DELAY



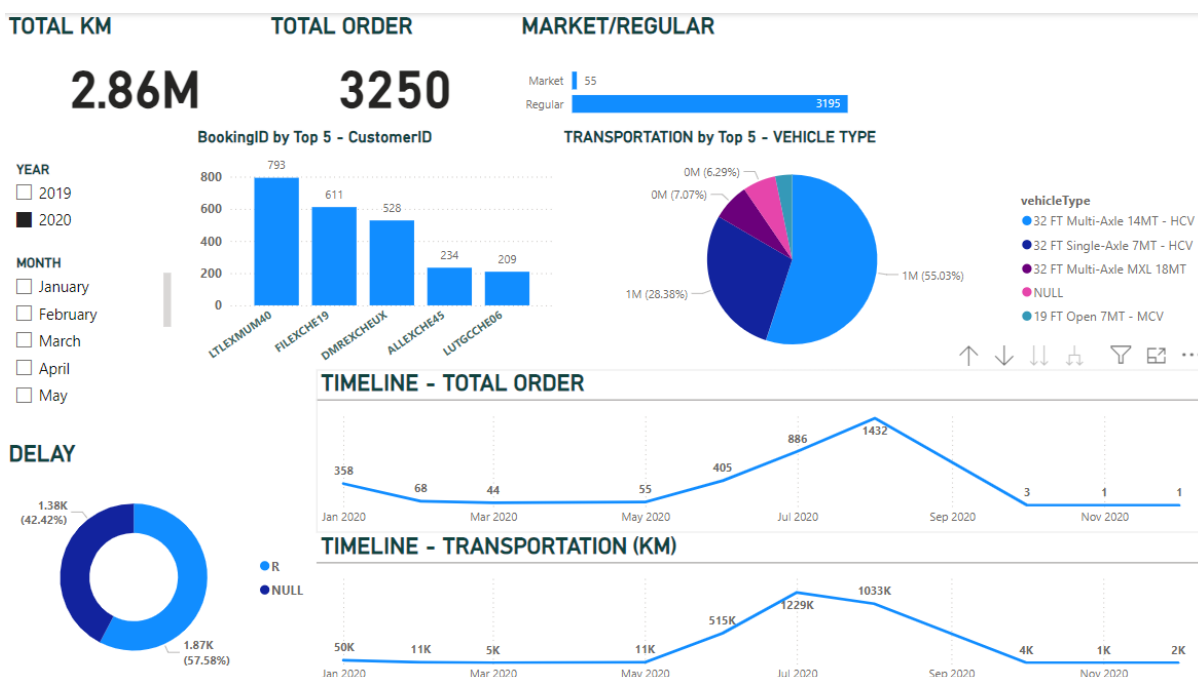
Hình 26: Tỷ lệ Delay 2019 - 2020

- **Phân tích giai đoạn năm 2019:**



Hình 27: Chỉ số tổng quan năm 2019

- Tổng Km đã đi là 351.82K Km.
  - Tổng số đơn hàng là 2720.
  - Đơn vị đặt hàng là LTLEXMUM40.
  - Tỷ lệ đơn bị trễ chiếm 81.11% tổng đơn hàng được đặt hàng.
  - Top 5 phương tiện được sử dụng nhiều nhất thì trong đó 40 FT 3XL Trailer 35MT, chiếm 47,65% tổng lượt sử dụng trong top 5.
  - Tháng có đơn hàng nhiều nhất là tháng 11 với 642 đơn hàng.
  - Tháng có tổng Km đi nhiều là tháng 12 với 83.042Km.
- **Phân tích giai đoạn năm 2020:**



Hình 28: Chỉ số tổng quan 2020

- Tổng Km đã đi là 2.86M Km. Tăng 8 lần so với năm 2019.
- Tổng đơn hàng là 3250. Tăng 19% so với năm 2019.
- Đã có nhiều đơn vị đặt hàng hơn so với năm 2019.
- Tỷ lệ đơn bị trễ chiếm giảm xuống còn 57.58% tổng đơn hàng được đặt so với 81.11% của năm 2019.
- Top 5 phương tiện được sử dụng nhiều nhất thì trong đó 32 FT Multi-Axle 14MT - HCV, chiếm 55.03% tổng lượt sử dụng trong top 5. Có sự thay đổi dòng xe so với năm 2019.
- Tháng có đơn hàng nhiều nhất là tháng 8 với 1.432 đơn hàng.
- Tháng có tổng Km đi nhiều là tháng 7 với 1.228.741Km.

### Nhận xét:

#### **Điểm mạnh:**

- Tăng trưởng mạnh mẽ về tổng quãng đường đã đi (tăng 8 lần).
- Sự gia tăng số đơn hàng và đơn vị đặt hàng cho thấy sự mở rộng thị trường.
- Cải thiện đáng kể tỷ lệ đơn hàng trễ (giảm từ 81.11% xuống còn 57.58%).
- Phương tiện chủ lực vẫn hoạt động ổn định và có sự thay đổi dòng xe để cải thiện hiệu suất.

#### **Điểm cần cải thiện:**

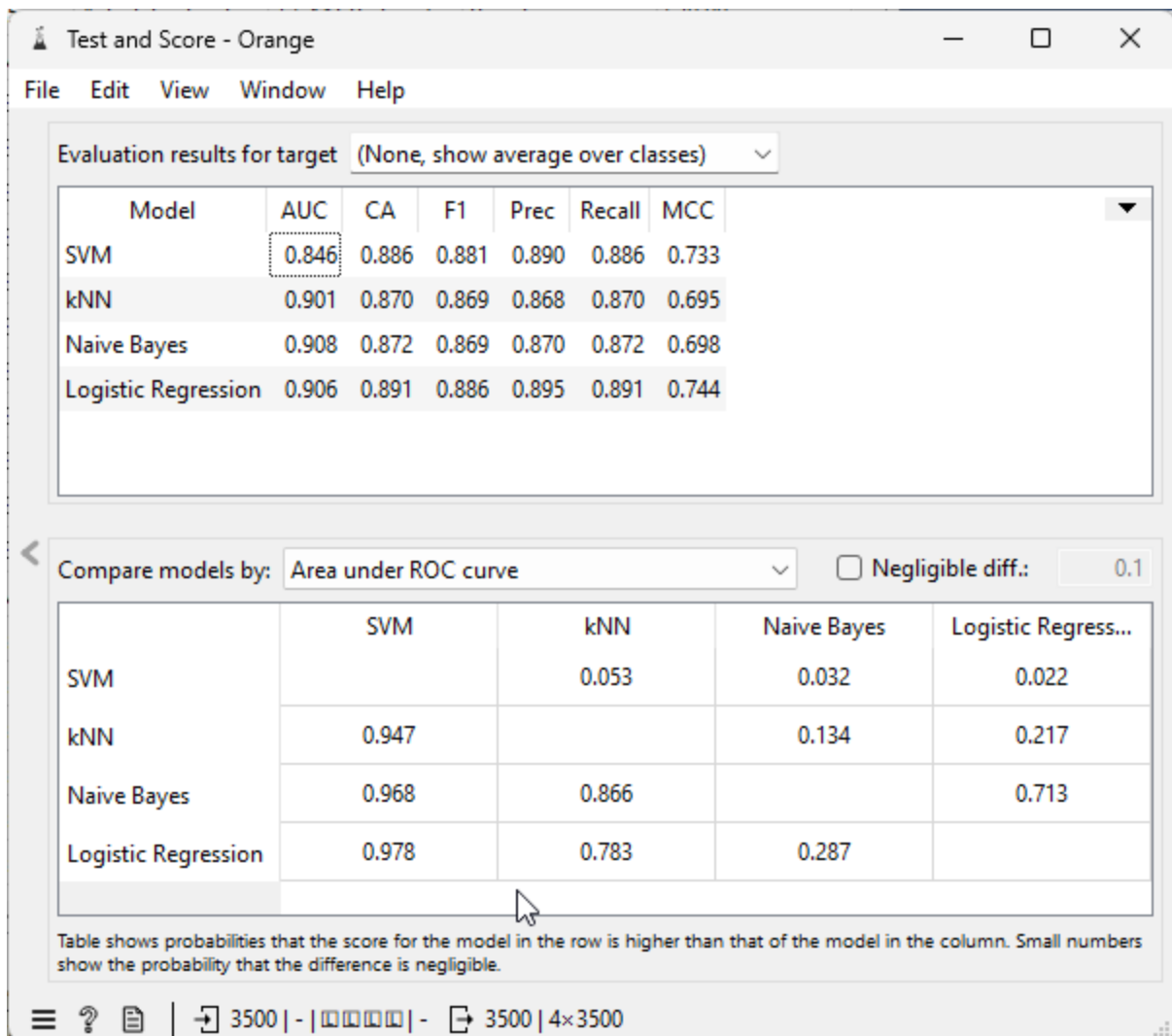
- Mặc dù số lượng đơn hàng tăng, nhưng tỷ lệ tăng trưởng về đơn hàng thấp hơn so với quãng đường đã đi, điều này có thể cần đánh giá lại việc tối ưu hóa các đơn hàng và chiến lược vận hành.
- Tỷ lệ đơn hàng trễ vẫn còn cao và cần tiếp tục giảm thiểu.
- Cần nghiên cứu thêm nguyên nhân tại sao tháng 8 có số lượng đơn hàng cao nhất và làm thế nào để duy trì hoặc tăng trưởng số đơn hàng trong các tháng khác.

## 8. Tổng hợp kết quả từ Orange

	Logistic Regression	supplierID	Material Shipped	supplierNameCod	BookingID	Planned ETA	trip_start_date	customerID	ustomerNameCoc	vehicleType	Market/Regular	RTATION_DISTANC
1	0.79 : 0.21 → G	73073	LU KUPPLUNGS...	Vijay Transport	AEIBK2026570	8/23/20 10:11 PM	8/20/2020 13:56	ALLEXCHE45	Ashok leyland li...	1 MT Tata Ace (...)	Regular	535.00
2	0.83 : 0.17 → G	67974	AUTO PARTS	S.M.G Transport	AEIBK2025944	8/18/20 2:50 PM	8/13/2020 20:00	ALLEXCHE45	Ashok leyland li...	1.5 MT Pickup (...)	Regular	530.00
3	0.89 : 0.11 → G	67974	EMPTY TRAYS	S.M.G Transport	AEIBK2025586	8/15/20 8:28 PM	8/11/2020 16:28	ALLEXCHE45	Ashok Leyland ...	1.5 MT Pickup (...)	Regular	530.00
4	0.41 : 0.59 → R	67824	AUTO PARTS	KIRAN BAJIRA...	AEIBK2024916	8/5/20 5:53 PM	8/5/2020 17:53	CSYEXCHEM2	Comstar autom...	1.5 MT Vehicle (...)	Regular	0.00
5	0.57 : 0.43 → G	68006	AUTO PARTS	ST TRANSPORT	AEIBK2022492	7/15/20 7:26 PM	7/11/2020 15:26	TMCGCCHE05	Tvs motor com...	1.5 MT Vehicle (...)	Regular	215.00
6	0.57 : 0.43 → G	68006	DAY QUARTER...	ST TRANSPORT	AEIBK2022654	7/17/20 6:37 PM	7/13/2020 14:37	TMCGCCHE05	Tvs motor com...	1.5 MT Vehicle (...)	Regular	215.00
7	0.63 : 0.37 → G	68004	EMPTY TRAYS	Sree Sakthi Tran...	AEIBK2022485	7/15/20 6:11 PM	7/11/2020 14:11	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
8	0.63 : 0.37 → G	68004	THERMAL CUT...	Sree Sakthi Tran...	AEIBK2022657	7/17/20 6:37 PM	7/13/2020 14:37	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
9	0.63 : 0.37 → G	68004	AUTO PARTS	Sree Sakthi Tran...	AEIBK2022442	7/14/20 8:09 PM	7/10/2020 16:09	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
10	0.63 : 0.37 → G	68004	LU WASHER FL...	Sree Sakthi Tran...	AEIBK2021839	7/6/20 6:42 PM	7/2/2020 14:42	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
11	0.63 : 0.37 → G	68004	AUTO PARTS	Sree Sakthi Tran...	AEIBK2023789	7/29/20 7:25 PM	7/25/2020 15:25	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
12	0.63 : 0.37 → G	68004	RECTIFIER ASSE...	Sree Sakthi Tran...	AEIBK2023876	7/31/20 3:18 PM	7/27/2020 11:18	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
13	0.63 : 0.37 → G	68004	AUTO PARTS	Sree Sakthi Tran...	AEIBK2021972	7/8/20 4:30 PM	7/4/2020 12:30	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
14	0.63 : 0.37 → G	68004	COP LEVER INN...	Sree Sakthi Tran...	AEIBK2022218	7/12/20 4:07 PM	7/8/2020 12:07	TMCGCCHE05	Tvs motor com...	15 FT Single-Ax...	Regular	215.00
15	0.48 : 0.52 → R	68004	SEALING / FUSS...	Sree Sakthi Tran...	AEIBK2025423	8/14/20 5:46 PM	8/10/2020 13:46	TMCGCCHE05	Tvs motor com...	17 FT Container	Regular	980.00
16	0.41 : 0.59 → R	67762	BEARING BUSH...	DAIWAN DEVID...	AEIBK2024657	8/3/20 12:02 PM	8/3/2020 10:56	CSYEXCHEM2	Comstar autom...	17 FT Open 5M...	Regular	47.00
17	0.87 : 0.13 → G	67789	STICKER / LATC...	GAYATHRI TRA...	AEIBK2025730	8/16/20 8:16 PM	8/12/2020 16:16	ALLEXCHE45	Ashok Leyland ...	17 FT Open 5M...	Regular	400.00
18	0.87 : 0.13 → G	67789	AUTO PARTS	GAYATHRI TRA...	AEIBK2027147	8/28/20 3:56 PM	8/21/2020 20:30	ALLEXCHE45	Ashok leyland li...	19 FT OPEN BO...	Regular	400.00
19	0.87 : 0.13 → G	67789	PLANETARY AXLE	GAYATHRI TRA...	AEIBK2027346	8/29/20 7:45 PM	8/26/2020 11:58	ALLEXCHE45	Ashok leyland li...	19 FT OPEN BO...	Regular	400.00
20	0.92 : 0.08 → G	67789	EMPTY FOAM	GAYATHRI TRA...	AEIBK2023922	7/31/20 7:44 PM	7/27/2020 15:44	ALLEXCHE45	Ashok Leyland ...	19 FT OPEN BO...	Regular	400.00
21	0.92 : 0.08 → G	2785	WIPER MOTOR ...	SREE MAHAA R...	AEIBK2024257	8/3/20 3:10 PM	7/30/2020 11:10	ALLEXCHE45	Ashok Leyland ...	19 FT OPEN BO...	Regular	400.00
22	0.92 : 0.08 → G	67789	SWITCH ASSY ...	GAYATHRI TRA...	AEIBK2023223	7/24/20 7:19 PM	7/20/2020 15:19	ALLEXCHE45	Ashok Leyland ...	19 FT OPEN BO...	Regular	400.00

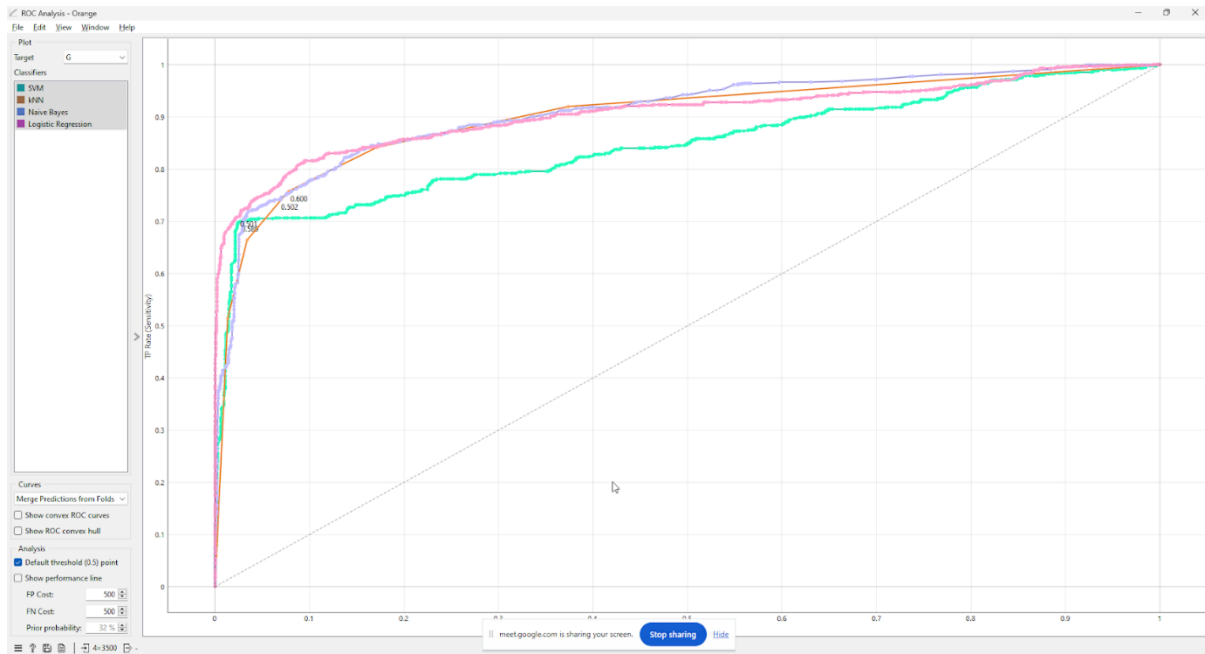
Hình 29: Mô hình huấn luyện Orange

Dựa vào Test and Score lựa chọn mô hình huấn luyện phù hợp.



Hình 30: Test and Score

Sử dụng ROC analysis để đánh giá.



Hình 31: ROC analysis

Các đường cong ROC của Naive Bayes và Logistic Regression nằm gần nhau và ở vị trí cao hơn so với SVM và kNN, cho thấy chúng có AUC cao hơn và do đó khả năng phân biệt tốt hơn.

- Knn

Learners		Predicted		
Actual		G	R	$\Sigma$
	G	82.0 %	10.9 %	1113
	R	18.0 %	89.1 %	2387
	$\Sigma$	1028	2472	3500

Hình 32: Kết quả kNN



- Naive Bayes

Learners		Predicted		
		G	R	$\Sigma$
Actual	G	83.7 %	11.5 %	1113
	R	16.3 %	88.5 %	2387
	$\Sigma$	986	2514	3500

Hình 33: Kết quả Naive Bayes

- Logistic Regression

Learners		Predicted		
		G	R	$\Sigma$
Actual	G	93.0 %	12.2 %	1113
	R	7.0 %	87.8 %	2387
	$\Sigma$	848	2652	3500

Hình 34: Kết quả Logistic Regression

Trong bối cảnh dự đoán chuyến hàng bị delay, cần ưu tiên giảm thiểu False Negative (FN) hơn là False Positive (FP). Điều này có nghĩa là nên tập trung vào việc xây dựng mô hình có độ phủ (Recall) cao, tức là khả năng phát hiện ra càng nhiều chuyến hàng bị delay càng tốt, ngay cả khi phải chấp nhận một số dự đoán sai (FP).

Vậy Logistic Regression là phương pháp có chỉ số tốt nhất cho cả 2 yếu tố giảm thiểu False Negative (FN) 7.0% (thấp nhất) và mô hình có độ phủ (Recall) cao 0.891 (cao nhất).