UNIVERSITÄT HOHENHEIM

Applied Data Science Lab (AOK Hackathon) WS24

AMBULANT-SENSITIVE HOSPITAL CASES

DATA-DRIVEN PREDICTION AND PROGNOSIS TOOL

Project Presentation, 23.01.2025

Team C

TEAM ORGANIZATION

M : Main Responsibility
S : Support

Tasks	Huyen	Grigory	Shengyan	Rita
1. Research kick-off information and study datasets + Descriptive analysis by answer task list's questions	M	M	M	M
2. Data cleaning and variables transformation	M			
3. Clustering Nursing Homes	S	M		
4. Variable Selection	M	M		
5. Model Selection	M	M	M	
6. Prognosis Tool	M			
7. Combine code files	S	S	M	
8. Prepare Presentation Slides:				
a. Introduction and Data Processing	M			
b. Descriptive Analysis				M
c. Clustering Nursing Homes & Variable Selection		M		
d. Model Selection			M	
e. Prognosis Tool – Use Cases	M			

CONTENT

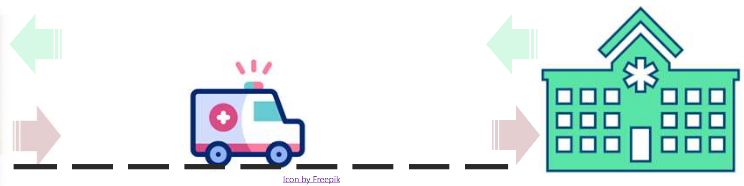
1. INTRODUCTION

- Summary situation and data generation process.
- Overview needs to reduce ambulance-sensitive cases.
- Objectives of Project.

1. INTRODUCTION
AMBULANT-SENSITIVE HOSPITAL CASES IN GERMANY

NURSING HOME

HOSPITAL



Variable *id_nursing_home* *id_case* **unique* ***Ambulance*** ***(Depend. Var.)*** *hospital_icd*

Nursing Home information

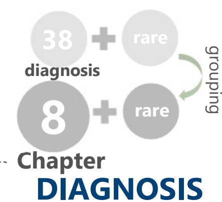
general information

HISTORY

	-Q4	-Q3	-Q2	-Q1	id_case
drug	Amlodipin			rare	0001
icd	I50	E11	rare	F06	0001

drug_history (-Q4 -Q3 -Q2 -Q1)

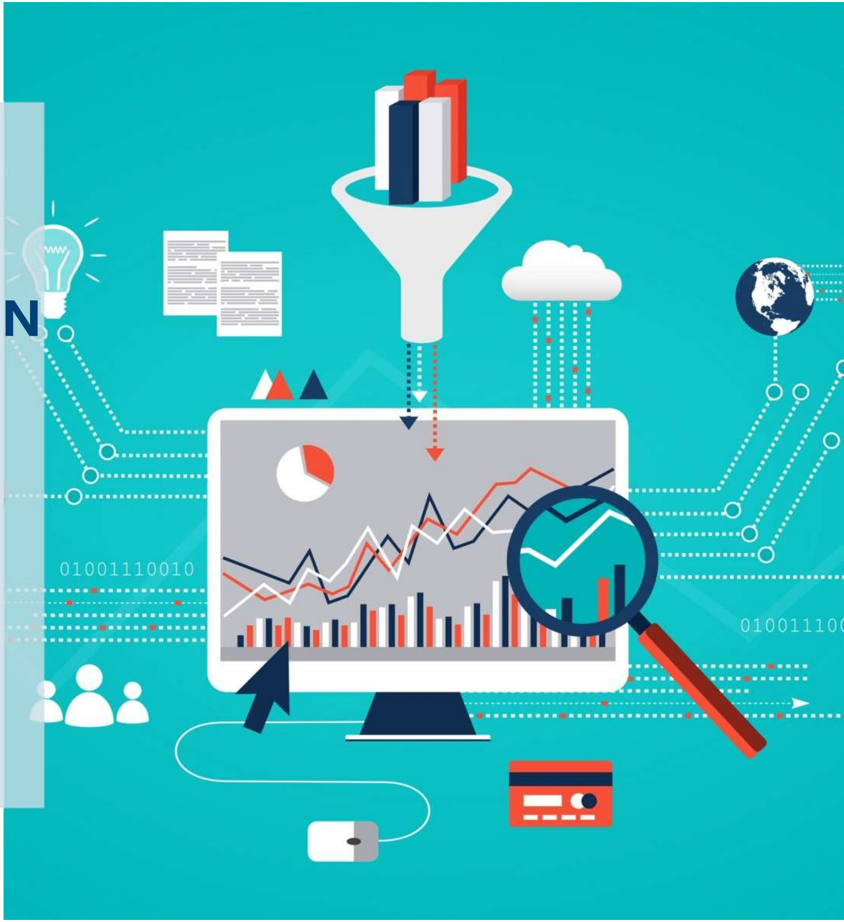
icd_history (-Q4 -Q3 -Q2 -Q1)



*Dtsch Arztebl 2015; 112(45): A-1860 / B-1538 / C-1498

2. DATA PROCESSING & VARIABLES CONSTRUCTION

- Create binary dependent variable 'ambulance'
- Transform drugs and diagnose historical text data into numeric and dummy variables
- Grouping ICD into Chapters to reduce high-dimension issues



2.1 SUMMARY DATA TABLE

id_case	hospital_icd	gender	age	care_level	id_nursing_home	ambulance	icd_name	icd_chapter	icd_chap_name
2012971112302230000	J44	m	80	4	1712709520361220000	1	Other chronic obstructive pulmonary disease	Chapter X	Diseases of the respiratory system
6373860270810270000	J44	m	60	2	-1578013449694770000	1	Other chronic obstructive pulmonary disease	Chapter X	Diseases of the respiratory system
-5314353177798480000	J44	m	60	2	-1578013449694770000	1	Other chronic obstructive pulmonary disease	Chapter X	Diseases of the respiratory system
4165098561186660000	I50	m	90	5	-6153414508441000000	1	Heart failure	Chapter IX	Diseases of the circulatory system
1077793242251920000	I50	m	70	3	-1646224910132520000	1	Heart failure	Chapter IX	Diseases of the circulatory system
6714796627062680000	J44	m	60	2	-1578013449694770000	1	Other chronic obstructive pulmonary disease	Chapter X	Diseases of the respiratory system
3352027258352980000	rare	m	70	3	3096284105978770000	1	Rare Diseases	Rare Diseases	Rare Diseases
181384588174676000	rare	m	70	3	3096284105978770000	1	Rare Diseases	Rare Diseases	Rare Diseases
-957566525004473000	I50	f	70	2	-191356676656247000	1	Heart failure	Chapter IX	Diseases of the circulatory system
227568413546568000	rare	m	90	3	-837439068155888000	1	Rare Diseases	Rare Diseases	Rare Diseases

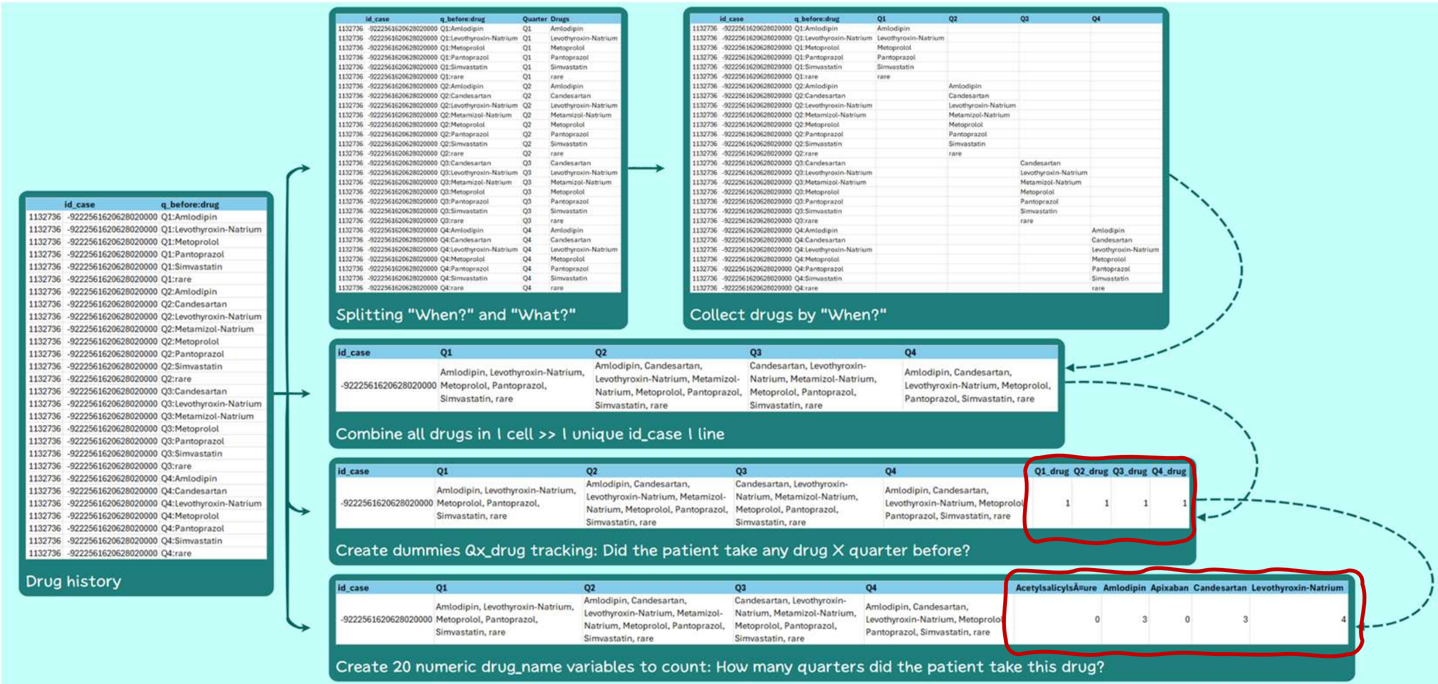
- ✓ **id_case:** unique for each patient visit, 1 patient can have multiple id_case but data is anonymous >> treated each case independently.
- ✓ **Id_nursing_home:** unique to each nursing home, a nursing home can have multiple patient visits. Lack of information of total residents per nursing home >> assumption scope reflected by total id_case
- ✓ **Control variables:** gender, age, care level

Dependent Variable:
'ambulance' – Dummy =1 if ambulance sensitive case, otherwise =0.


Icon by Freepik

>>Target of prediction model

2.2 DRUG HISTORY DATA TABLE



2.3 DIAGNOSE HISTORY DATA TABLE

id_case		hospital_icd	gender	age	care_level	id_nursing_home	ambulance	icd_name	icd_chapter	icd_chap_name	Q1_drug	Q2_drug	Q3_drug	Q4_drug	AcetylsalicylsÃ¤ure	Amlodipin	Candesartan	rare	Q1_icd	Q2_icd	Q3_icd	Q4_icd	Chapter IV	Chapter IX	Chapter XIV	Rare Diseases
1132736	-9222561620628020000	I50	f	70	4	-8211795106232080000	1	Heart failure	Chapter IX	Diseases of the circulatory system	1	1	1	1	0	3	3	4	1	1	1	1	4	4	4	4

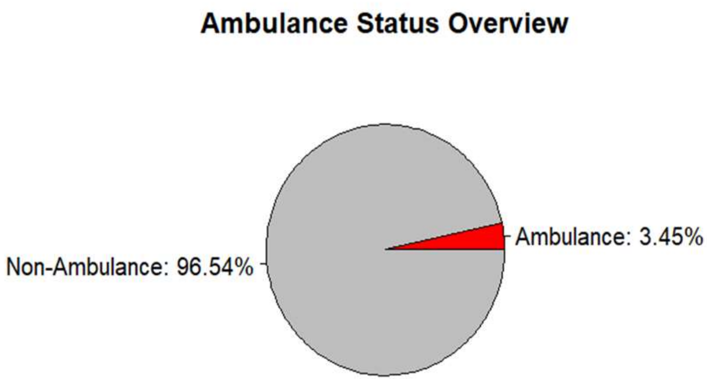
- ✓ **ICD Codes Grouping:** ICD codes were grouped into Chapters based on the official icd_head_code list from WHO.
- ✓ **Transformation of History Variables:** All history variables stored in text were transformed into numeric variables to reflect the frequency or duration of drugs/diagnoses in patients' medical history.
- ✓ **QX_drug & QX_icd Variables:** These are dummy variables tracking the specific quarter when the medical history was recorded.
- ✓ **drug_history and icd_history variables:** sum of 4 QX_drug & QX_icd dummy variables [0;4]
- ✓ **<u>Potential Trade-Offs:</u>** Transforming data from text to numeric may lead to some loss of information, but it enhances compatibility for statistical models and ensures consistency across the dataset.

3. DESCRIPTIVE ANALYSIS

- Describe the important items in a statistical manner.
- Which diagnoses most often lead to hospital admissions?
- Are there differences in patients with or without admission?
- Are there differences between different nursing homes?

DESCRIPTIVE STATISTICS OF KEY VARIABLES

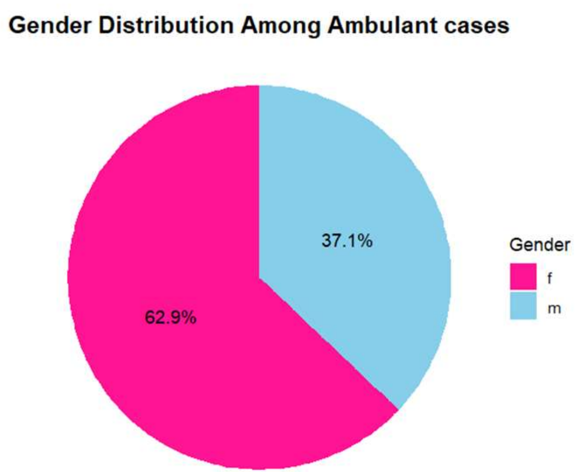
Ambulance Rate



- Total ambulant patients: 10431
- Number of ambulant patients without any prior ICD diagnosis: 21
- Percentage of ambulant patients without prior ICD diagnosis: 0.2 %

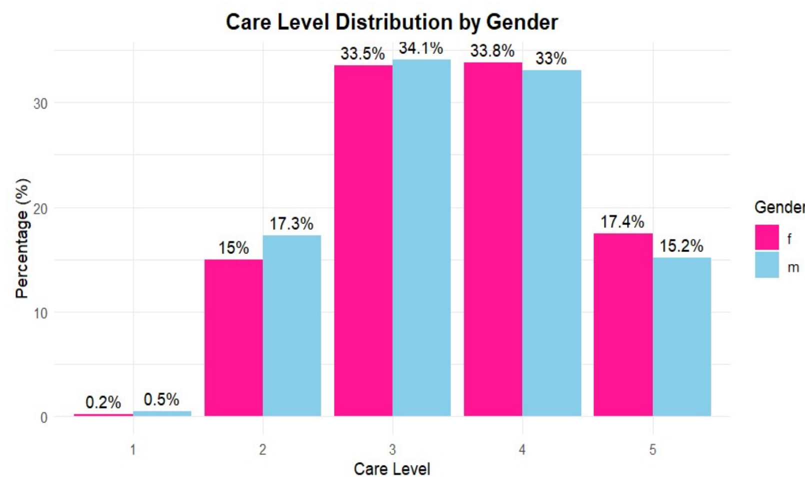
Gender proportions for ambulant cases

- Among ambulant cases, females accounted for the majority, representing 62.9%, while males comprised 37.1%. This indicates a higher ambulance rate for females compared to males.



Are Male or Female Patients Associated with Different Care Levels?

- Females and males are almost evenly distributed, with slightly more females in Care Level 3 (33.5% for females vs. 34.1% for males) and more males in Care Level 4 (33.8% for females vs. 33% for males)
- Care Level 1 is the least common for both genders, with very low percentages (0.2% for females and 0.5% for males).
- In Care Level 2, there are slightly more males and care level 5 has more female



The most used drugs

- The table displays the most commonly prescribed drugs.
- The drug Rare stands out as the most prescribed, accounting for 14.9% of patients.

Drug	Count
rare	6477
Metamizol-Natrium	4271
Torasemid	3961
Pantoprazol	3777
Ramipril	2492
Amlodipin	1893
Metoprolol	1865
Bisoprolol	1828
Levothyroxin-Natrium	1805
Acetylsalicylsäure	1793

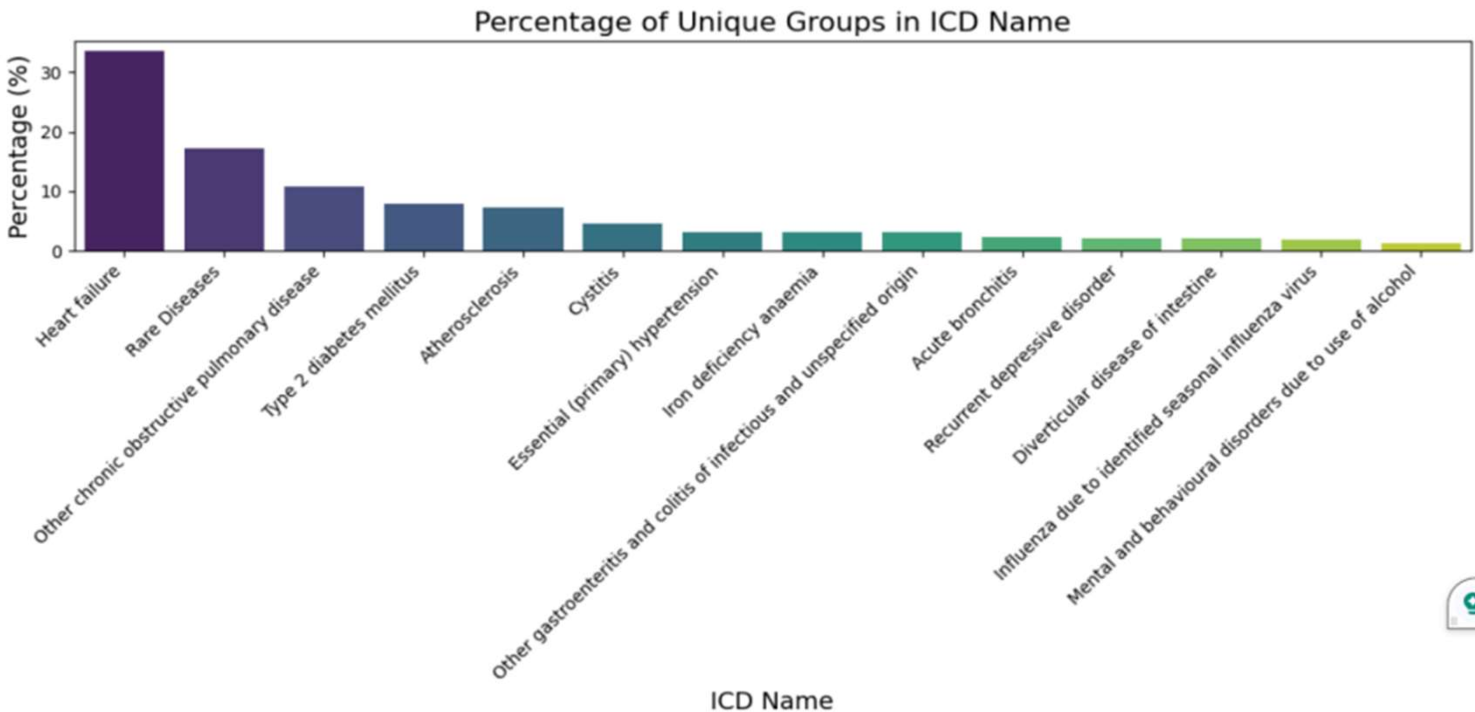
Descriptive Statistics of Diagnoses per Patient

This table summarizes the distribution of the number of diagnoses per ambulant patient:

Mean	Median	Min	Max
14.7	13	1	39

TOP DIAGNOSES LEADING TO HOSPITAL ADMISSION

By ICD name



TOP DIAGNOSES LEADING TO HOSPITAL ADMISSION

By chapter of ICD

Diagnosis Chapter	Count
Diseases of the circulatory system	4572
Rare Diseases	1789
Diseases of the respiratory system	1567
Endocrine, nutritional and metabolic diseases	815
Diseases of the genitourinary system	476
Mental and behavioural disorders	353
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	321
Certain infectious and parasitic diseases	316
Diseases of the digestive system	222

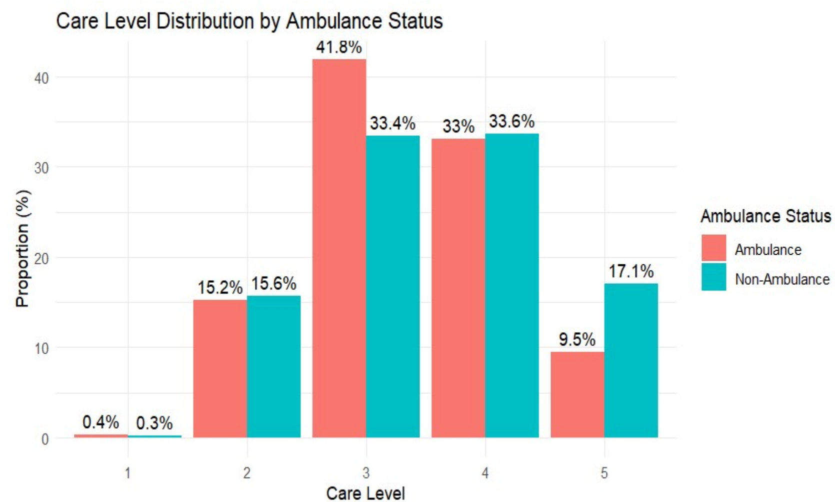
DIFFERENCES IN PATIENTS WITH AND WITHOUT ADMISSIOM

Proportion of cases by care level

- Care Level 3 has the highest proportion of ambulant patients (41.8%), making it the most common among admissions.
- Care Levels 4 and 5 have similar distributions between ambulance and non-ambulance cases.
- Care Levels 1 and 2 are the least represented overall.

Mean and median age for ambulance and non-ambulance cases

Ambulance status	Count	Mean age	Median age
Ambulance	10,431	82.57789	80
Non-ambulance	291,190	84.00721	80



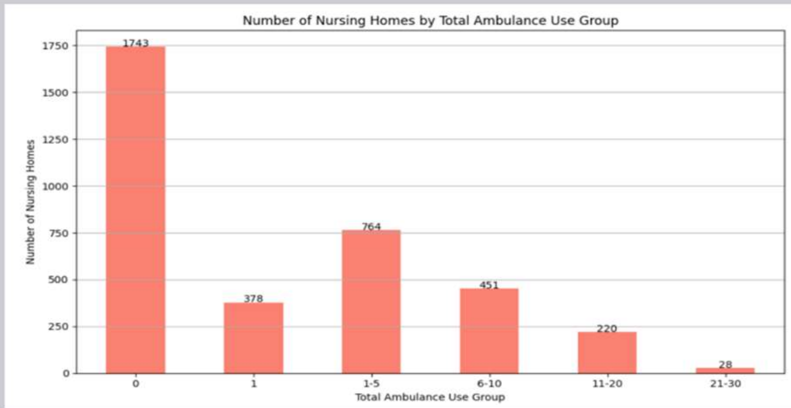
Frequently prescribed drugs between ambulance vs. non-ambulance

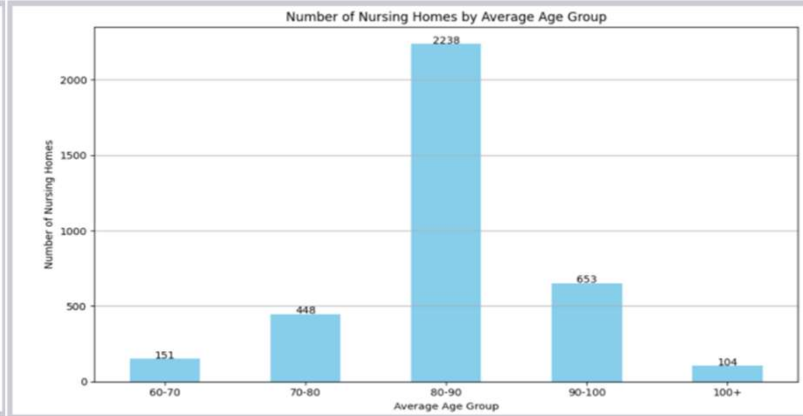
- Ambulant patients are prescribed significantly more drugs, averaging **5.55 drugs per patient**, compared to just **0.034 drugs** for non-ambulant patients.
- The drug **Rare** is the most commonly prescribed, accounting for **54.2% of ambulance cases**.
- Patients taking a higher number of drugs are at greater risk of ambulance.

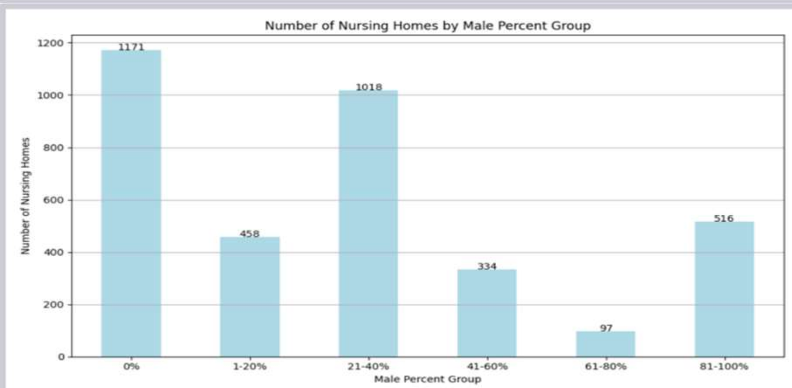
Ambulance status	Mean drugs	Median drugs
Ambulance	5.55	6
Non-ambulance	0.0340	0

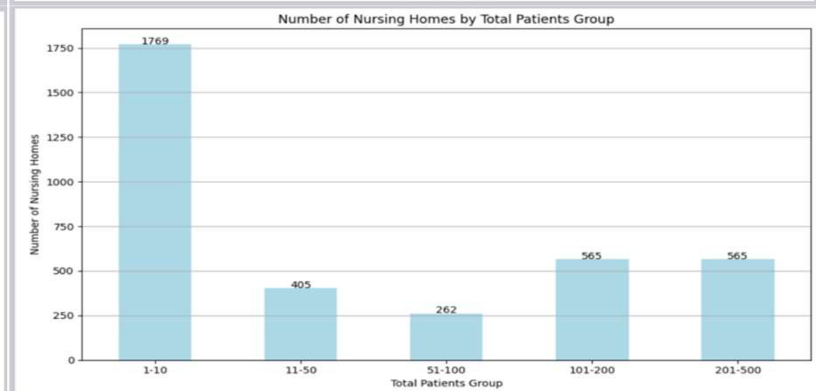
Non-ambulance		
Drug	**Count**	**Percentage**
Rare	5656	54.2%
Metamizol-Natrium	4145	39.7%
Torasemid	3875	37.1%
Pantoprazol	3657	35.1%
Ramipril	2394	23.0%

EXPLORING DIFFERENCES BETWEEN NURSING HOMES

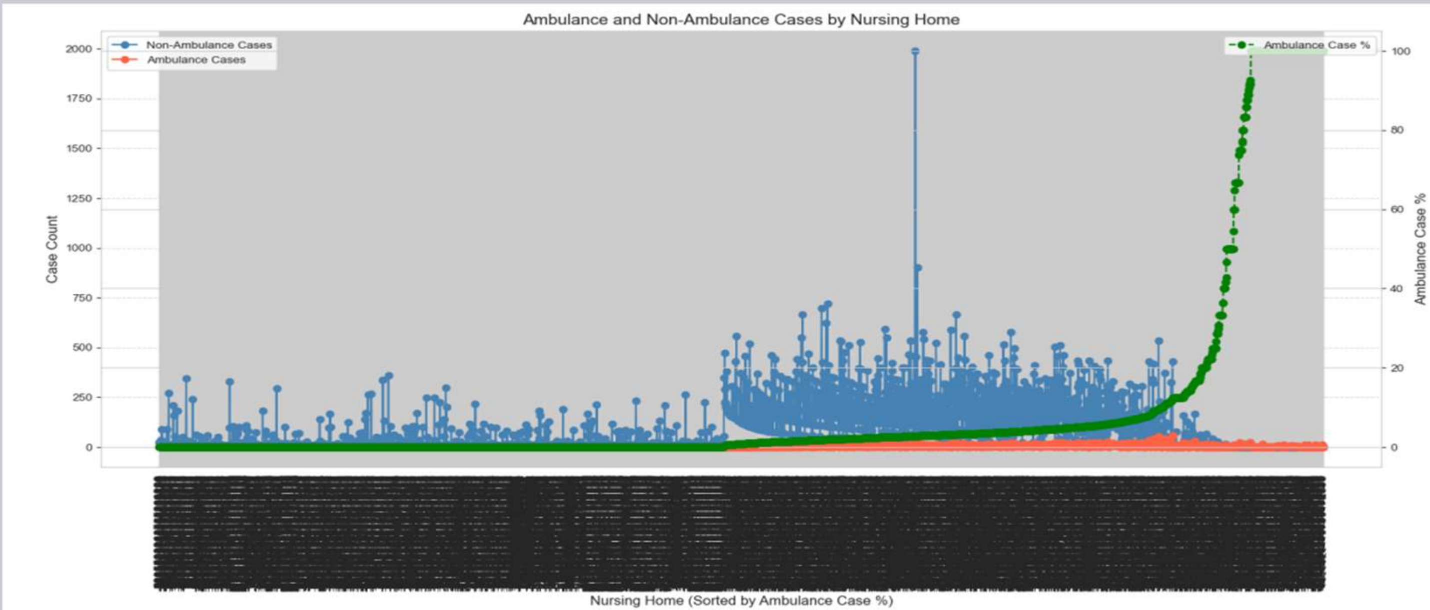








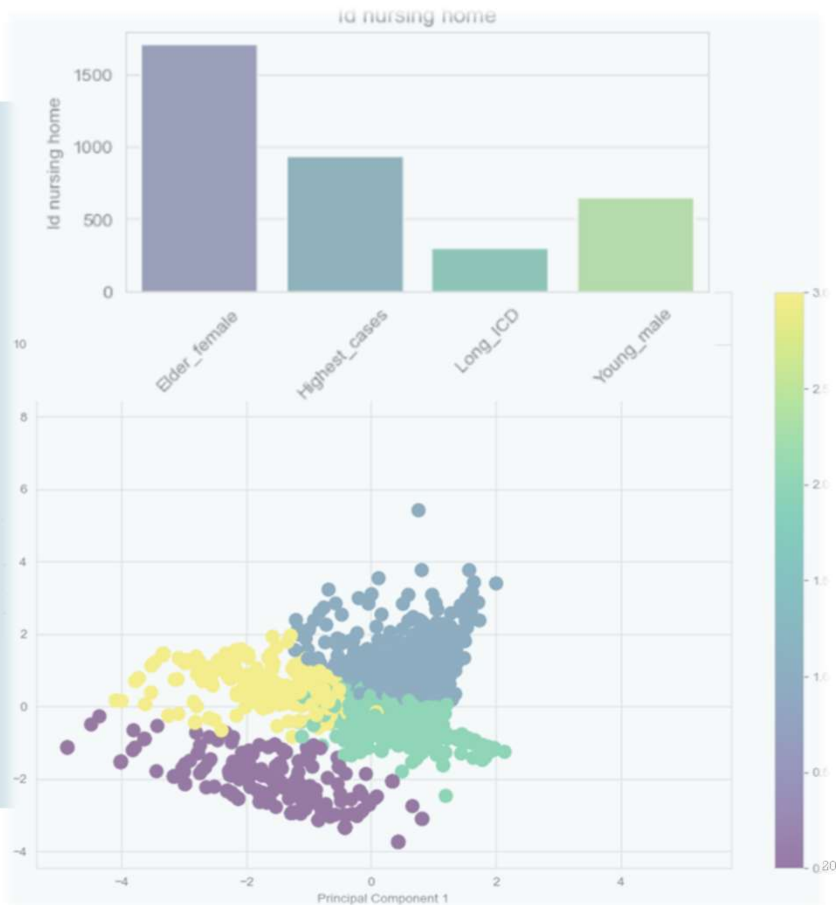
EXPLORING DIFFERENCES BETWEEN NURSING HOMES



Distribution of ambulance cases by nursing homes

4. CLUSTERING NURSING HOMES

- Use the Elbow method to identify optimal K of K-mean clustering.
- Clustering nursing homes into 4 groups based on 4 features: age, gender, scope and long diagnose history of patients.
- Compare the final groups to get the group's definition.



FIND OPTIMAL K BY ELBOW METHOD:

The **K-Means elbow method** - determine the optimal number of clusters

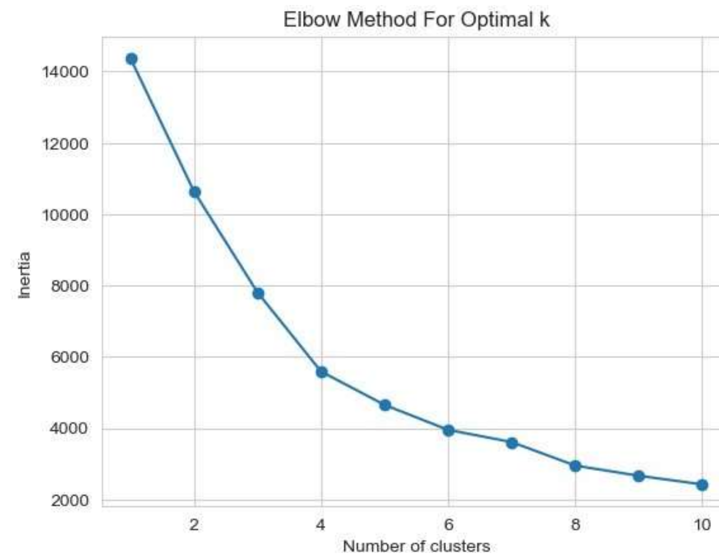
slowly decrease ->good trade-off between the **number of clusters** and the **amount of variance explained**

Average Age

ICD History

Female Percent

Total Patients



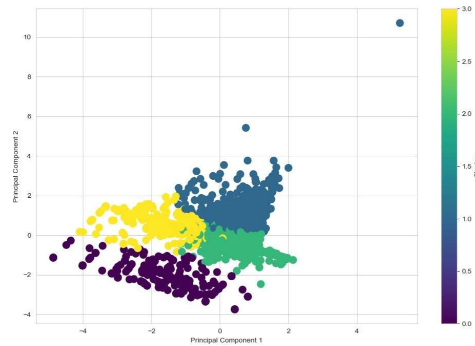
CLUSTER GROUP ELDER FEMALE

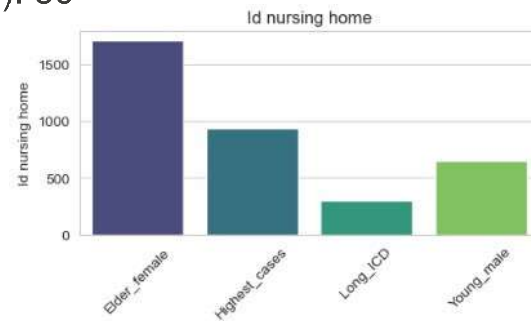
Old female in small nursing homes with small amount of diseases

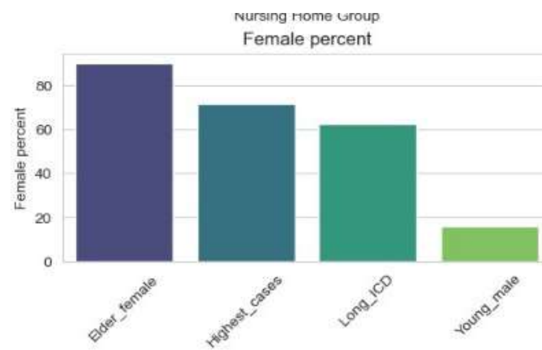
Average Age: 85.59 years

Female Percent: 89.81%

Total Patients (mean): 30







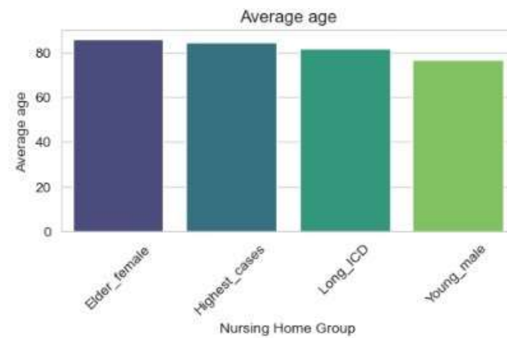
CLUSTER GROUP YOUNG MALE

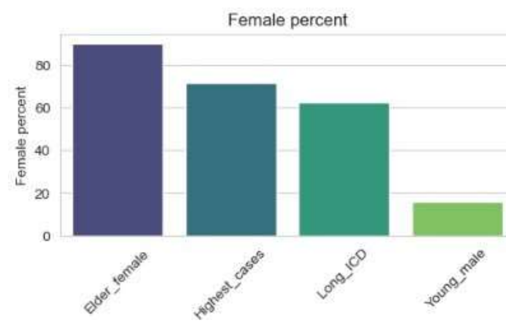
Group with young male

Total Patients (mean): 17

Average age: 76.5

Short icd history





CLUSTER GROUP HIGHEST CASES

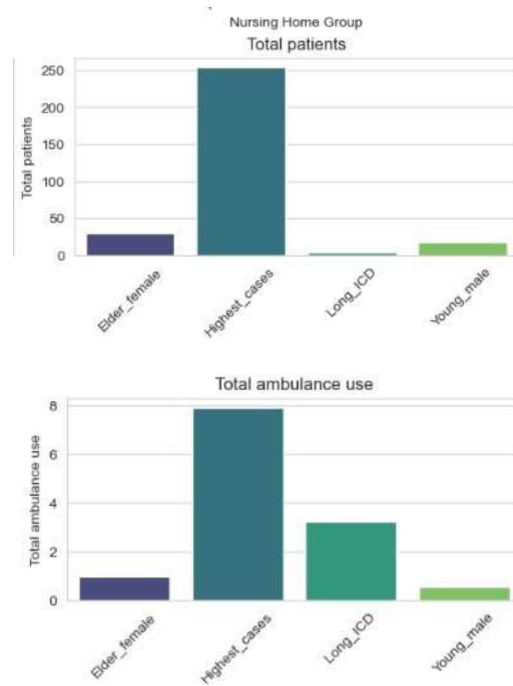
Group with highest number of ambulant cases pro nursing home

Biggest nursing homes

Total Patients (mean): 253

Ambulance Use Percent: 3.12%

Absolute ambulant case number: 7.9



CLUSTER GROUP LONG DISEASE

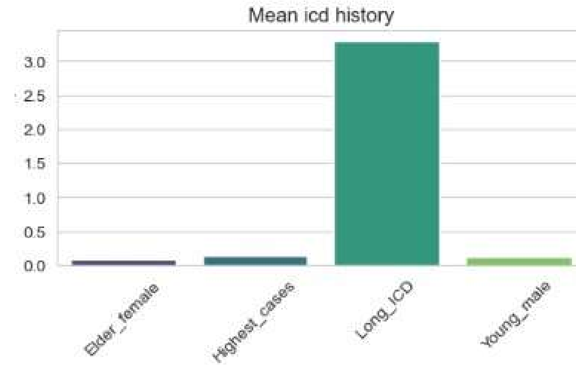
Group with highest percentage of ambulant cases pro nursing home

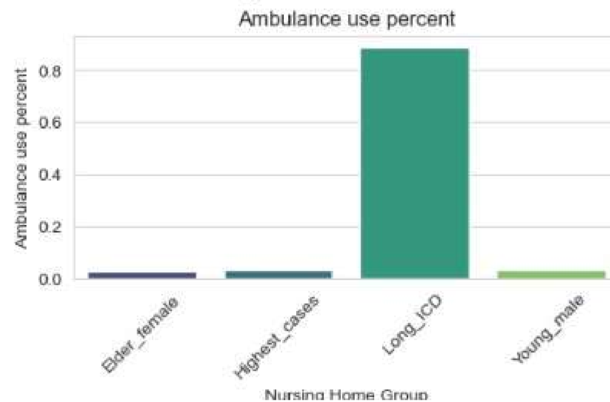
Ambulance Use Percent: 88.63%

Patients with long disease history

Total Patients (mean):4

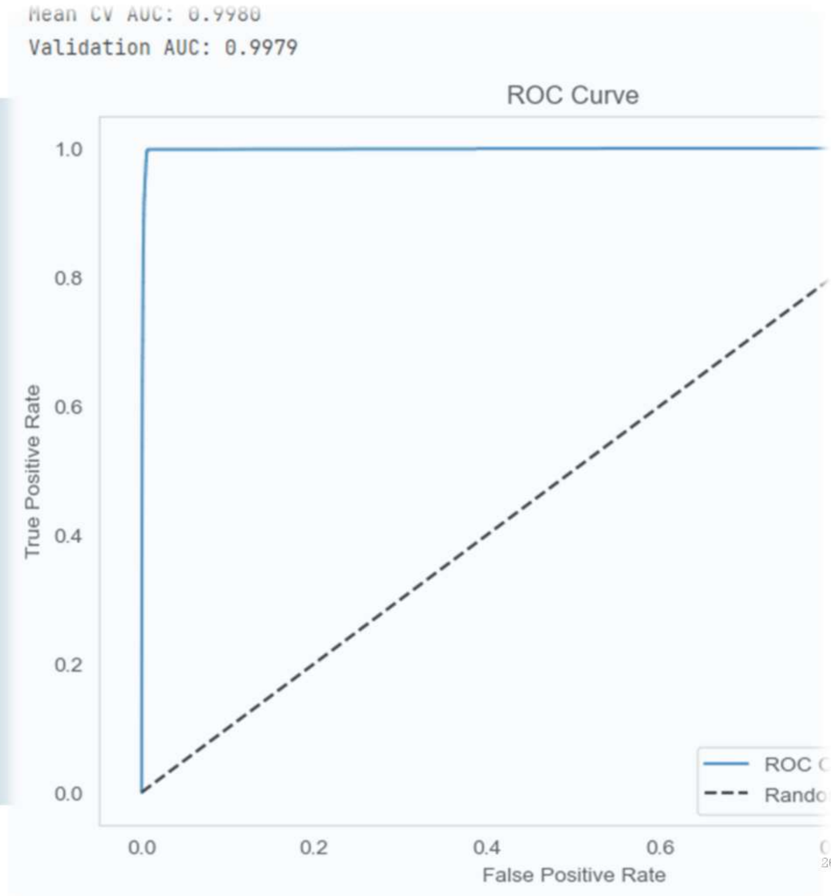
Absolute ambulant case number: 3.2





5. VARIABLE SELECTION

- Used logistic regression to test the variable groups' important
- Used XGBoost to rank the variable important for final variable subset.
- Finalize the variable subset for the model selection.



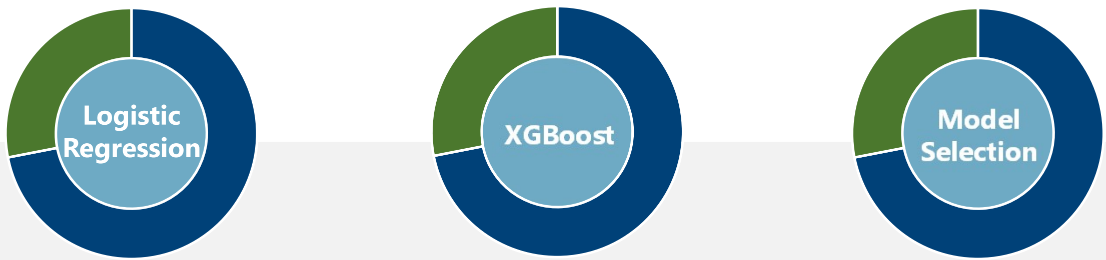
WHY VARIABLE SELECTION

Process of identifying and selecting a subset of relevant features

1. Improves Model Performance (reduces the chance of Overfitting)
2. Enhances Interpretability (helps to grasp the factors influencing outcome)
3. reduces the computational cost

VARIABLE SELECTION PROCESS

1. Separate all variables to independent subsets
2. Run reduced models (XGBoost and Logistic Regression)
3. Chose top of important Features from each subset
4. Combine chosen features in a Full model.
5. Chose the most important features



Test variable groups important

6 subsets: Drugs, ICD, Chapters, Nursing Home Groups, Long History, Quarter History.

Ranking and choosing variables

3 subsets: Drugs, ICD, Chapters.

Selected variable subset

Final variables: general variables, top_10 Drug history, top_10 Chapter history.

LOGISTIC REGRESSION

1.6 Subsets – reduced models

2.top 10 most important Variables

3.Model with all important variables

4.Chose the most important Features with Full model (all relevant features)

The coefficient tells how much the odds (or probability) of the outcome (like getting the disease) change when you increase a feature by **one unit**.

Subset model with drug history

	Feature	Coefficient
	Candesartan	5.800644
	Torasemid	4.423667
	Q1_drug	3.584156
	Metamizol-Natrium	3.410174
	Metoprolol	3.007920

25	age	-0.117206
21	Tilidin und Naloxon	-0.169256
1	Q2_drug	-0.393677
24	gender	-0.418286
26	care_level	-0.919536
3	Q4_drug	-1.083258

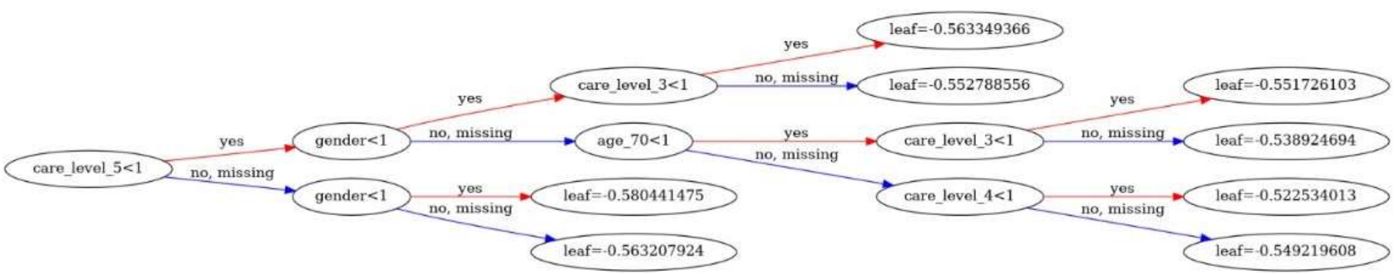
XGBOOST

-only with icd history (grouped to chapters)

-only drug history-only icd history (grouped by diseases)1.run 3 Models independently

2.Chose 10 Features with biggest F-score from each

```
Feature Importances:
care_level: 216.0
age: 189.0
Chapter XXI: 188.0
Chapter V: 164.0
Chapter XVIII: 150.0
nursing_home_group: 136.0
Chapter IV: 132.0
Chapter XIV: 131.0
Metamizol-Natrium: 105.0
Chapter XIII: 104.0
```



6. MODEL SELECTION

- Train 3 different models based on same datasets: train, validation and test data
- Compare models' performance by using AUC of the ROC curve
- Balance between accuracy and interpretation

MODEL SELECTION

- **Variables in our models:**

```
# Define variable groups
general_var = ['nursing_home_group', 'age', 'gender', 'icd_history', 'drug_history', 'care_level']
t10_chapters_quarter = ['Chapter IV', 'Chapter IX', 'Chapter XIV', 'Chapter XVIII',
                        'Chapter XXI', 'Rare Diseases', 'Chapter XIII', 'Chapter V',
                        'Chapter VI', 'Q3_icd']
t10_drug_quarter = ["Amlodipin", "Bisoprolol", "Metamizol-Natrium", "Metoprolol",
                    "Pantoprazol", "Ramipril", "Risperidon", "Torasemid", "rare_drugs", 'Q3_drug']
dependent_Y = ['ambulance']
```

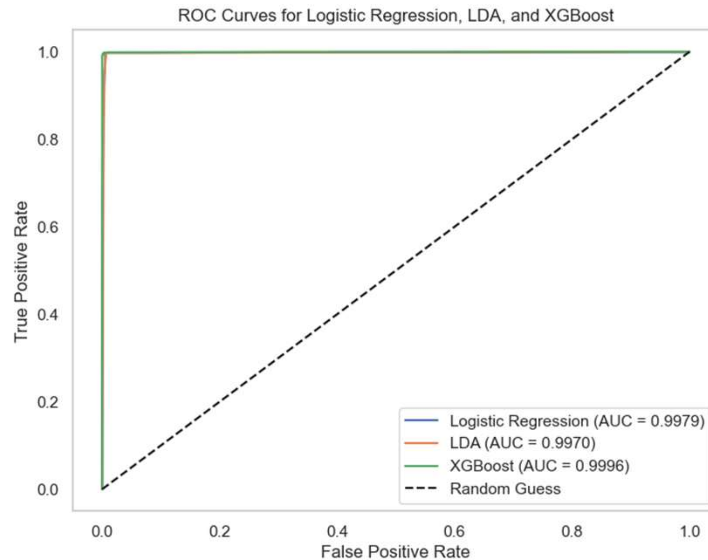
MODEL SELECTION

Models	Advantages	Challenges
XGBoost	Excels at **non-linear** relationships Can achieve **high performance** with proper tuning	**Complex** hyperparameter tuning (e.g., learning rate, max depth); **Interpretation** is harder than linear models; Overfitting risk without **constraints** (e.g., early stopping)
LDA (Linear Discriminant Analysis)	**Fast** to train, **low** computational cost Shows **class separation** in a linear subspace	**Normality** assumption; unequal covariances hurt performance; Hard to interpret with **many variables;** Large feature sets can cause **overfitting**
Logistic Regression	**Interpretable** (coefficients → log-odds impact) **Less** prone to overfitting vs. very complex models **Simple** to implement and tune	**Linear** decision boundary only; May underfit for **highly non-linear** data; Coefficients can be **unstable** without regularization if many features

COMPARE MODELS' PERFORMANCE

- 3 models had similar performance; all AUC are above 99%.
- Based on the clear differentiation of the data by nursing home groups, the models demonstrate very high performance.
- We applied 5-fold cross-validation and lasso regulation in the Logistic regression model to avoid overfitting.

```
Logistic Regression AUC: 0.9979
LDA AUC: 0.9970
XGBoost AUC: 0.9996
```



MODEL SELECTION

- **Logistic Regression** ---------Why We Chose Logistic Regression

Strong Cross-Validation Performance: Achieved accuracy and AUC scores comparable to XGBoost without complex hyperparameter tuning.

Simplicity & Interpretability: Coefficients are directly interpretable, which is crucial for explaining predictions in our case.

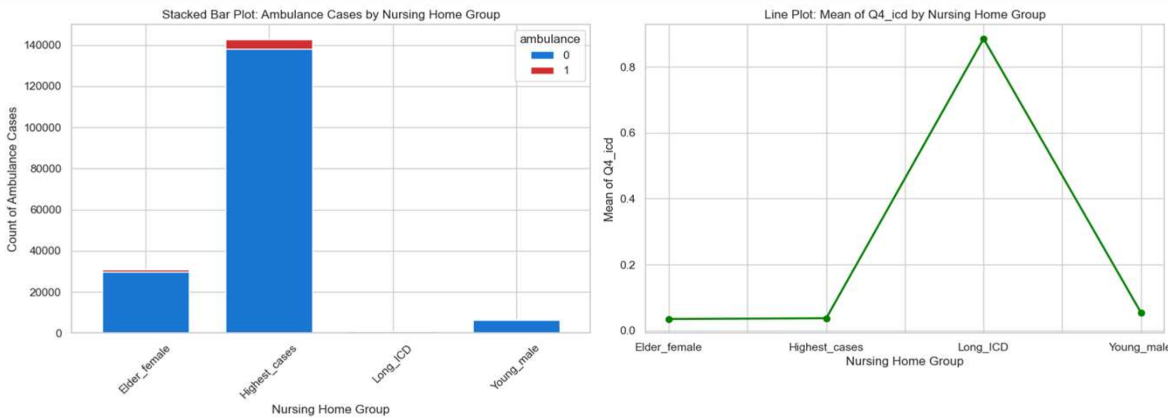
Robustness: Less risk of overfitting compared to more flexible models like XGBoost, especially on relatively modest datasets.

Logistic Regression gave us a good balance of prediction quality, interpretability, and ease of implementation. Hence, we selected it as our final model after confirming consistent performance via 5-fold cross-validation.

7. PROGNOSIS TOOL

- Train Logistic Regression as the final model for prediction tool
- Build a GUI option to check by case or full cases of 1 nursing home
- Use cases example
- Conclusions

7.1 FINAL PREDICTION MODEL

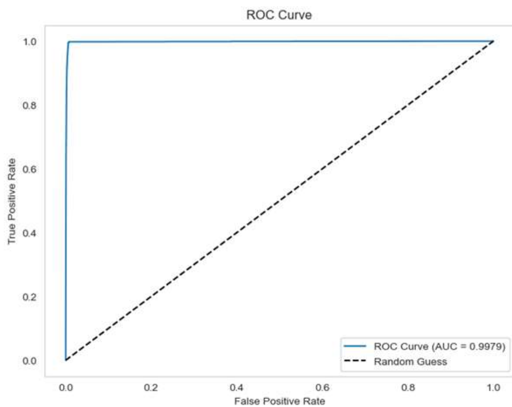


- The overall distribution of the training dataset across Nursing Home groups shows that the majority of ambulance cases are concentrated in the Highest_cases group.
- However, the Long_ICD group stands out with a significantly higher percentage of ambulance cases relative to the total cases within that group.
- This highlights the distinctiveness of the Long_ICD group in terms of its ambulance case density.

7.1 FINAL PREDICTION MODEL

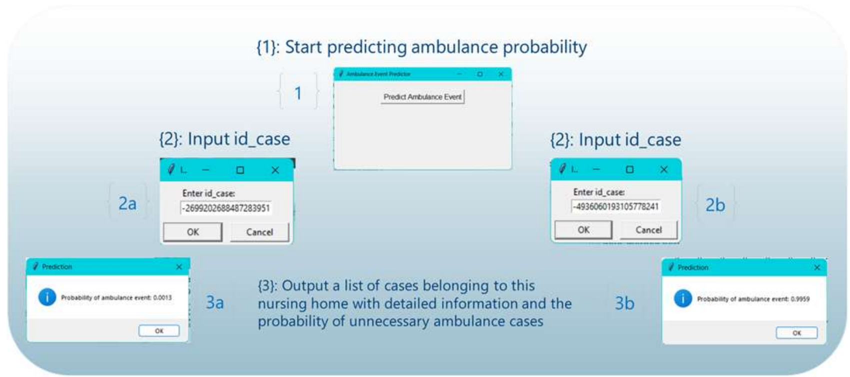
```
# Define variable groups
general_var = ['nursing_home_group', 'age', 'gender', 'icd_history', 'drug_history', 'care_level']
t10_chapters_quarter = ['Chapter IV', 'Chapter IX', 'Chapter XIV', 'Chapter XVIII',
                        'Chapter XXI', 'Rare Diseases', 'Chapter XIII', 'Chapter V',
                        'Chapter VI', 'Q3_icd']
t10_drug_quarter = ["Amlodipin", "Bisoprolol", "Metamizol-Natrium", "Metoprolol",
                    "Pantoprazol", "Ramipril", "Risperidon", "Torasemid", "rare_drugs", 'Q3_drug']
dependent_Y = ['ambulance']
```

```
Mean CV AUC: 0.9980
Validation AUC: 0.9979
```



Top 10 Variables Based on Coefficients:

	Variable	Coefficient
3	icd_history	2.782412
11	Rare Diseases	-1.757747
4	drug_history	-0.514092
23	Torasemid	0.357170
10	Chapter XXI	0.324636
7	Chapter IX	0.313511
20	Pantoprazol	0.230512
6	Chapter IV	0.221005
5	care_level	-0.200636
18	Metamizol-Natrium	0.183175

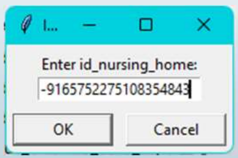


7.2 NURSING HOME USE CASES

{1}: Start screening nursing home

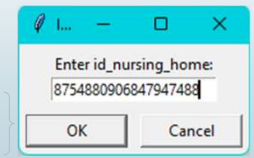
{2}: Input id_nursing_home

2a

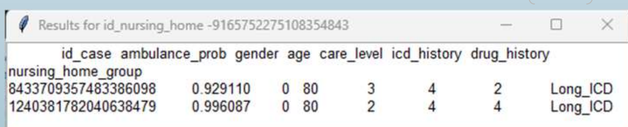


{2}: Input id_nursing_home

2b

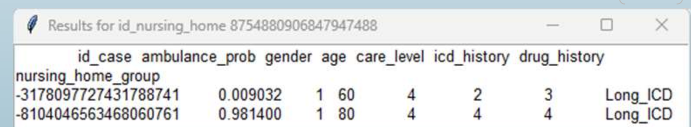


3a


Results for id_nursing_home -9165752275108354843

id_case	ambulance_prob	gender	age	care_level	icd_history	drug_history	nursing_home_group
8433709357483386098	0.929110	0	80	3	4	2	Long_ICD
1240381782040638479	0.996087	0	80	2	4	4	Long_ICD

3b


Results for id_nursing_home 8754880906847947488

id_case	ambulance_prob	gender	age	care_level	icd_history	drug_history	nursing_home_group
-3178097727431788741	0.009032	1	60	4	2	3	Long_ICD
-8104046563468060761	0.981400	1	80	4	4	4	Long_ICD

{3}: Output a list of cases belonging to this nursing home with detailed information and the probability of unnecessary ambulance cases

7.2 NURSING HOME USE CASES

Examples screening nursing home

Results for id_nursing_home -9210274116072207264

id_case	ambulance_prob	gender	age	care_level	icd_history	drug_history	nursing_home_group	ambulance
-7478417699530781837	0.001613	1	70	3	0	0	Highest_cases	0
759671494398263197	0.001337	1	90	3	0	0	Highest_cases	0
6368416783111750906	0.995472	0	90	3	4	0	Highest_cases	1
5154127692633154135	0.001469	1	80	3	0	0	Highest_cases	0
-771083166695072863	0.000988	1	100	4	0	0	Highest_cases	0
2334688295079119349	0.998889	1	70	2	4	2	Highest_cases	1
-7894120215296857176	0.001707	0	80	3	0	0	Highest_cases	0
4681651313393713847	0.001649	1	90	2	0	0	Highest_cases	0
3456058623770685175	0.001191	1	80	4	0	0	Highest_cases	0
-1686667024051285566	0.001337	1	90	3	0	0	Highest_cases	0
-5342305077085621039	0.001085	1	90	4	0	0	Highest_cases	0
3939389311026714803	0.002104	0	80	2	0	0	Highest_cases	0
6860453057908480264	0.001191	1	80	4	0	0	Highest_cases	0
7105486889718997578	0.001385	0	80	4	0	0	Highest_cases	0
8530014642666131234	0.001337	1	90	3	0	0	Highest_cases	0
700882549545607 43	0.001218	1	100	3	0	0	Highest_cases	0
6719213021382578982	0.001261	0	90	4	0	0	Highest_cases	0
3753195938983321943	0.001649	1	90	2	0	0	Highest_cases	0
6790999115955242863	0.001218	1	100	3	0	0	Highest_cases	0
-3249918001031982164	0.001191	1	80	4	0	0	Highest_cases	0
545549740692138967	0.001085	1	90	4	0	0	Highest_cases	0

Results for id_nursing_home -9141602204768789714

id_case	ambulance_prob	gender	age	care_level	icd_history	drug_history	nursing_home_group	ambulance
4460091267853444011	0.001337	1	90	3	0	0	Highest_cases	0
831374543427695519	0.001308	1	70	4	0	0	Highest_cases	0
-4682873706888854463	0.001308	1	70	4	0	0	Highest_cases	0
-9167126641053789837	0.002058	0	60	3	0	0	Highest_cases	0
1808626630016319870	0.001771	1	60	3	0	0	Highest_cases	0
8517386174633454445	0.001085	1	90	4	0	0	Highest_cases	0
8492963610936465716	0.001023	0	90	5	0	0	Highest_cases	0
3297288322020339634	0.001061	1	70	5	0	0	Highest_cases	0
-6214374971538226443	0.001649	1	90	2	0	0	Highest_cases	0
2771028615294528242	0.001337	1	90	3	0	0	Highest_cases	0
6036312601328903961	0.015075	1	70	5	3	3	Highest_cases	0
-5627361705279447959	0.001337	1	90	3	0	0	Highest_cases	0
-2584403544594995557	0.001771	1	60	3	0	0	Highest_cases	0
4675978233518603606	0.001085	1	90	4	0	0	Highest_cases	0
6671992280607633887	0.001085	1	90	4	0	0	Highest_cases	0
3624512062699433029	0.001085	1	90	4	0	0	Highest_cases	0
-3498964907052577142	0.001649	1	90	2	0	0	Highest_cases	0
-8171682138726558052	0.001061	1	70	5	0	0	Highest_cases	0

7.2 NURSING HOME USE CASES

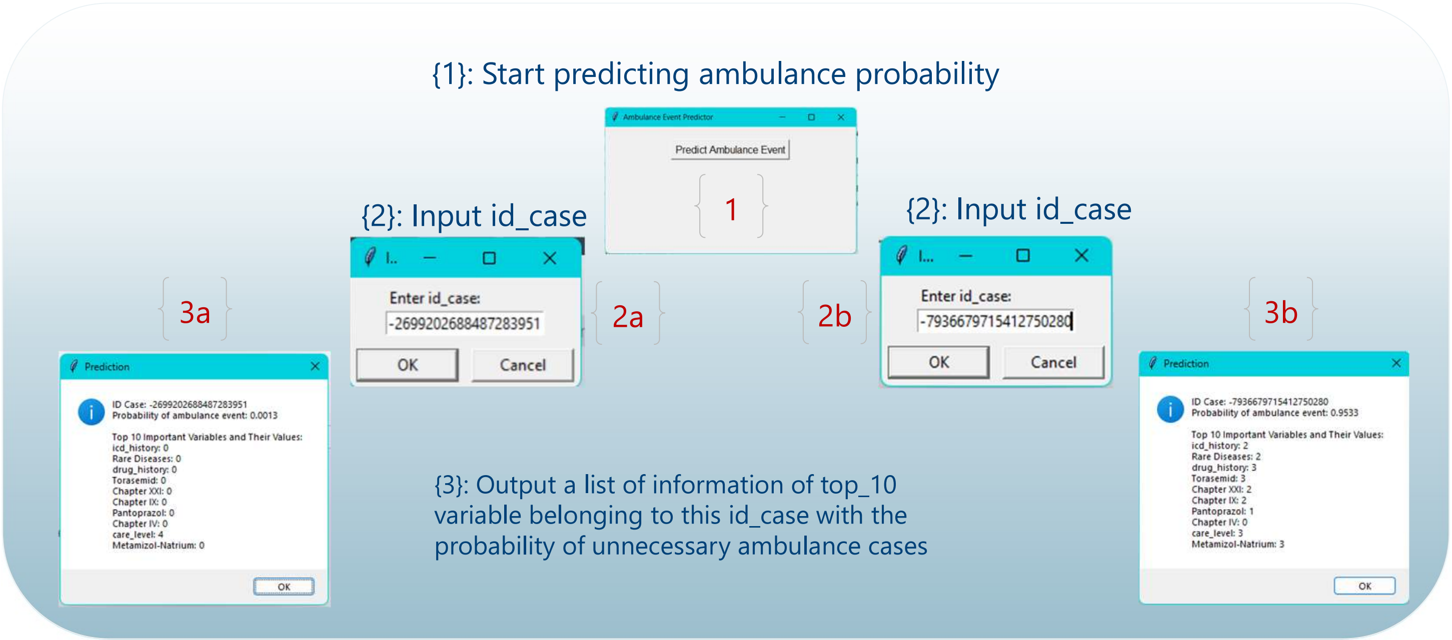
Examples screening nursing home

Results for id_nursing_home -9210274116072207264

id_case	ambulance_prob	ambulance	icd_history	Rare Diseases	drug_history	Torasemid	Chapter XXI	Chapter IX	Pantoprazol	Chapter IV	care_level	Metamizol-Natrium
-7478417699530781837	0.001613	0	0	0	0	0	0	0	0	0	3	0
759671494398263197	0.001337	0	0	0	0	0	0	0	0	0	3	0
6368416783111750906	0.995472	1	4	4	0	0	3	4	0	3	3	0
5154127692633154135	0.001469	0	0	0	0	0	0	0	0	0	3	0
-771083166695072863	0.000988	0	0	0	0	0	0	0	0	0	4	0
2334688295079119349	0.998889	1	4	4	2	2	2	4	2	0	2	2
-7894120215296857176	0.001707	0	0	0	0	0	0	0	0	0	3	0
4681651313393713847	0.001649	0	0	0	0	0	0	0	0	0	2	0
3456058623770685175	0.001191	0	0	0	0	0	0	0	0	0	4	0
-1686667024051285566	0.001337	0	0	0	0	0	0	0	0	0	3	0
-5342305077085621039	0.001085	0	0	0	0	0	0	0	0	0	4	0
3939389311026714803	0.002104	0	0	0	0	0	0	0	0	0	2	0
6860453057908480264	0.001191	0	0	0	0	0	0	0	0	0	4	0
7105486889718997578	0.001385	0	0	0	0	0	0	0	0	0	4	0
8530014642666131234	0.001337	0	0	0	0	0	0	0	0	0	3	0
700882549545607 43	0.001218	0	0	0	0	0	0	0	0	0	3	0
6719213021382578982	0.001261	0	0	0	0	0	0	0	0	0	4	0
3753195938983321943	0.001649	0	0	0	0	0	0	0	0	0	2	0
6790999115955242863	0.001218	0	0	0	0	0	0	0	0	0	3	0
-3249918001031982164	0.001191	0	0	0	0	0	0	0	0	0	4	0
545549740692138967	0.001085	0	0	0	0	0	0	0	0	0	4	0

- ✓ The Prognosis Tool can screen all the cases belong to 1 nursing home and run the prediction model to return the probabilities of ambulance sensitive case.
- ✓ Along with probabilities, the tool also can extract the value of top_10 most important variables for each case.
- ✓ Based on this nursing home's doctors can consider to take some prevention actions.

7.3 INDIVIDUAL USE CASES



THANK YOU

Applied Data Science Lab (AOK Hackathon) WS24