Cardiff Business School

Ysgol Busnes Caerdydd

ISF 2022 Oxford England

Forecasting length of stay in trauma network

Presenter(PhD students):

Zihao Wang WangZ121@cardiff.ac.uk

Supervision teams:

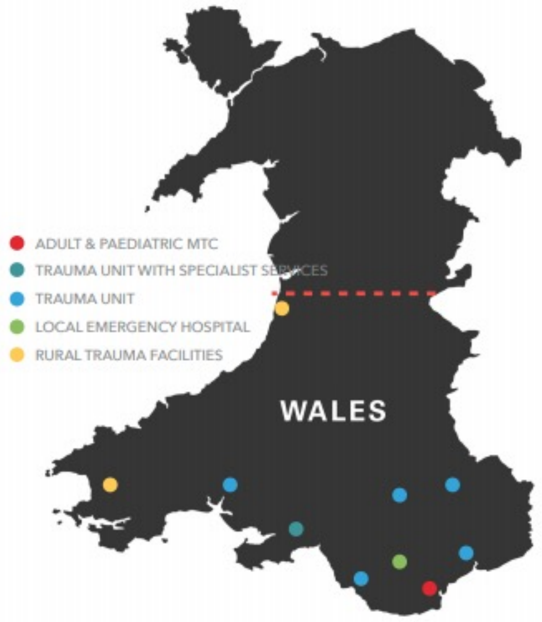
Bahman Rostami-Tabar Rostami-TabarB@cardiff.ac.uk

Jane Haider Haider@cardiff.ac.uk

Mohamed Naim NaimMM@cardiff.ac.uk

CONTENTS

Introduction of trauma networks



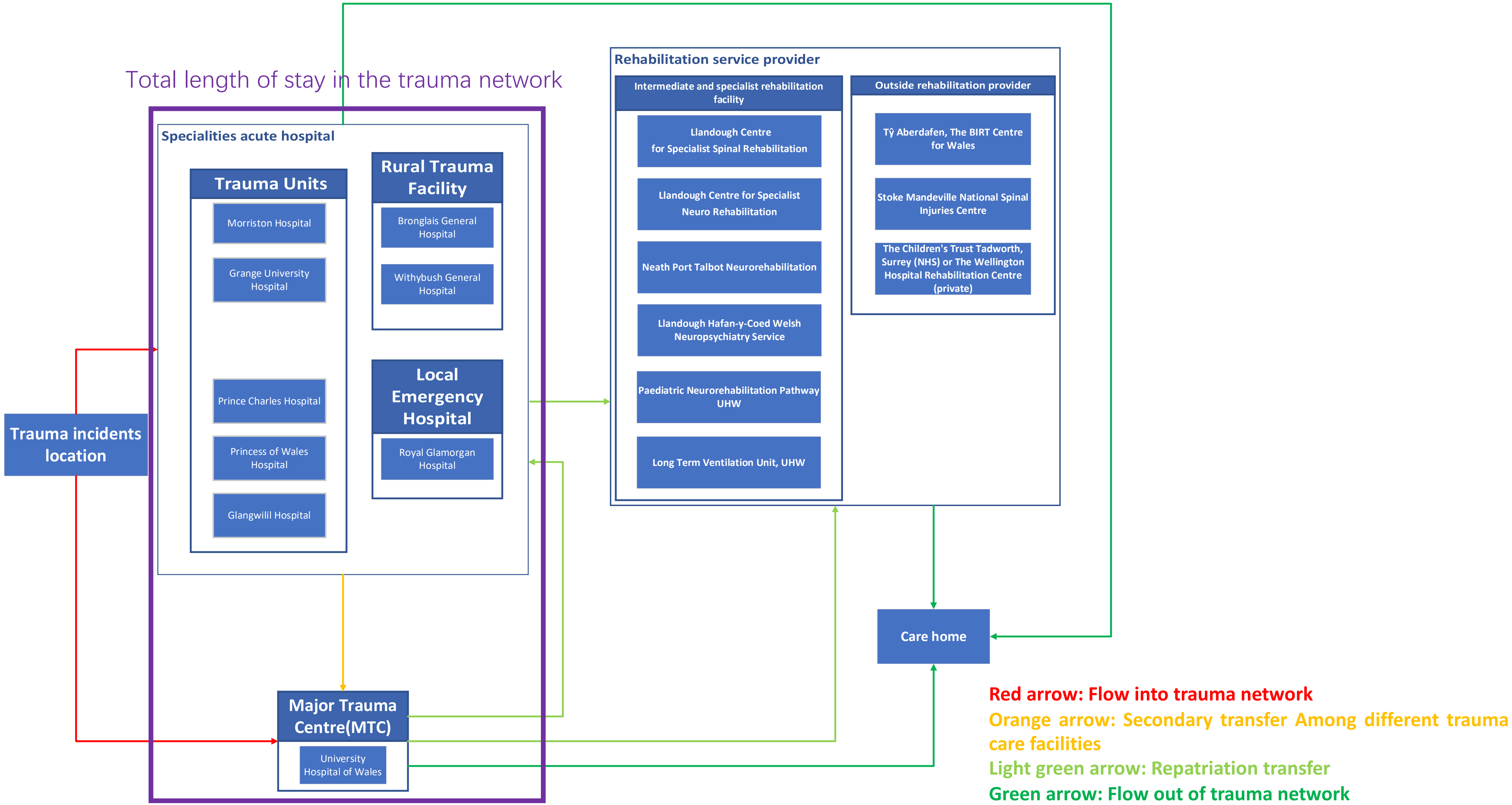
What is Trauma networks?

The collaboration between the health care providers to deliver trauma care services in a geographical area

Why networks?

- Organizing trauma care has been a strong health care agenda in UK since the 2010.
- Trauma networks in London and other place of England have shown a significant reduction in mortality/ morbidity and improvements in functional outcome.

General patient flow in trauma network

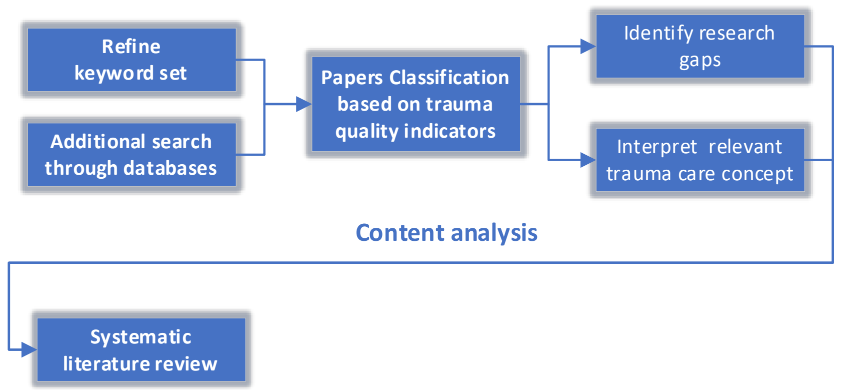


Previous work and Research objective

Phase 1: Systematic literature review

Aims:

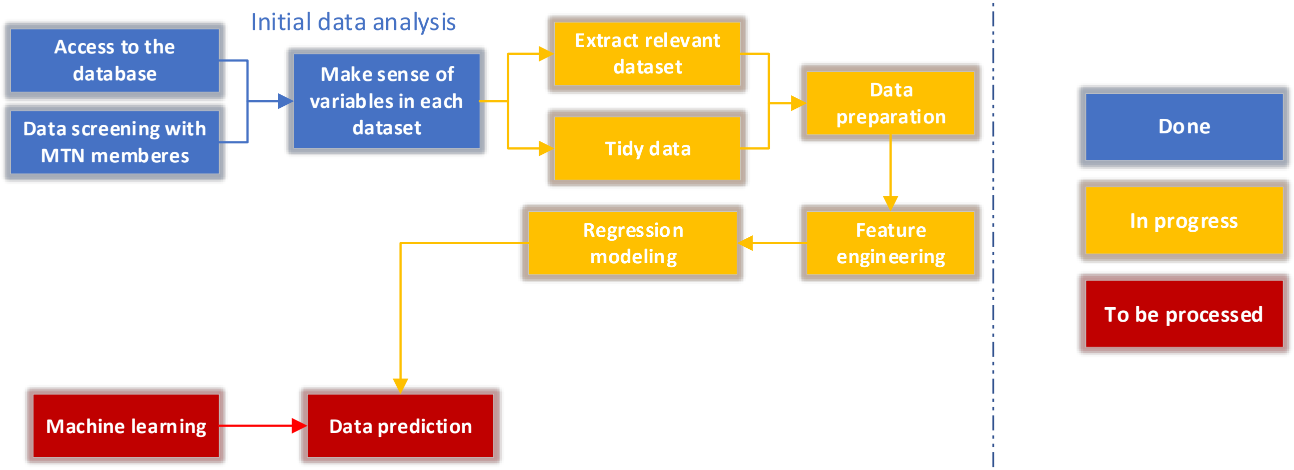
- **Identify and categories potential quality indicators for evaluating trauma network operational performance**
- **Understand general operation process in the trauma network**



Phase 2: Length of stay forecasting

Aims:

Explore the association between length of stay and other predictors and fit-in a effective model for length of stay forecasting



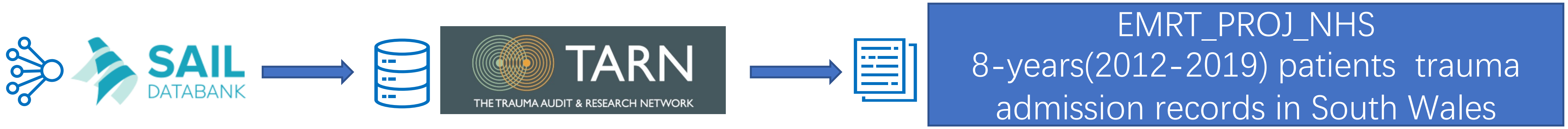
- Patient length of stay as a process indicator for evaluating the operation of trauma networks has rarely been explored.
- the clinical trauma interventions, patients' complications and discharge destination contribute to the variation of patients' length of stay within the trauma network (Moore 2018 et al.)

Long-term objective: the data-driven LOS forecasting model could provide analytical assumption to build the system dynamic model for simulating the patient flow in the whole trauma network

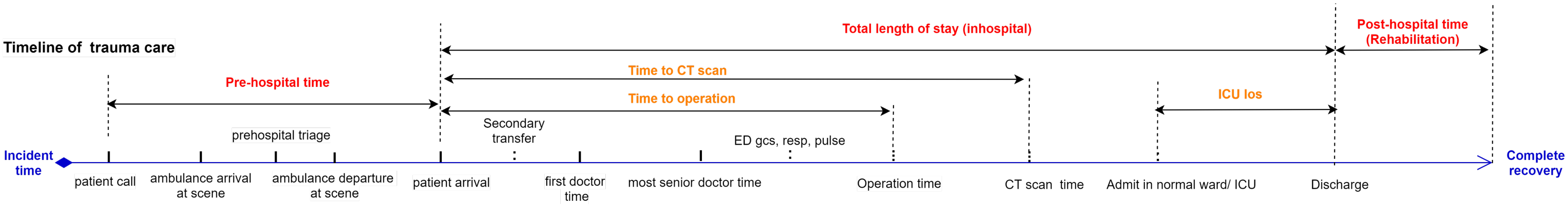
CONTENTS

Data acquisition and characteristics

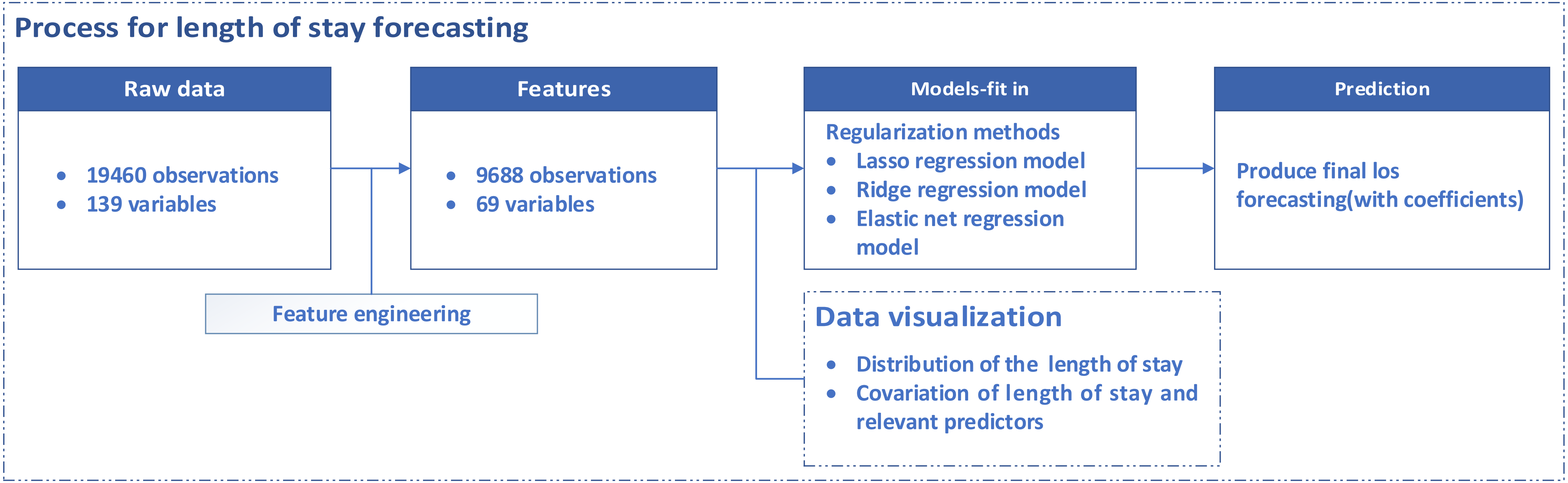
- Data Source



- Response variable： Total length of stay
- Predictors:
  - Categorical: age, gender, injury mechanism, transfer status, ward and Binary variables (Welsh incident, MTC,Operation…),
  - Numerical: ISS, ED&prehospital GCS, RESP, SBP, Pulse…
  - Time series: time points and interval in the trauma care process



Data analysis process



Feature engineering (Process and Challenge)

Feature engineering

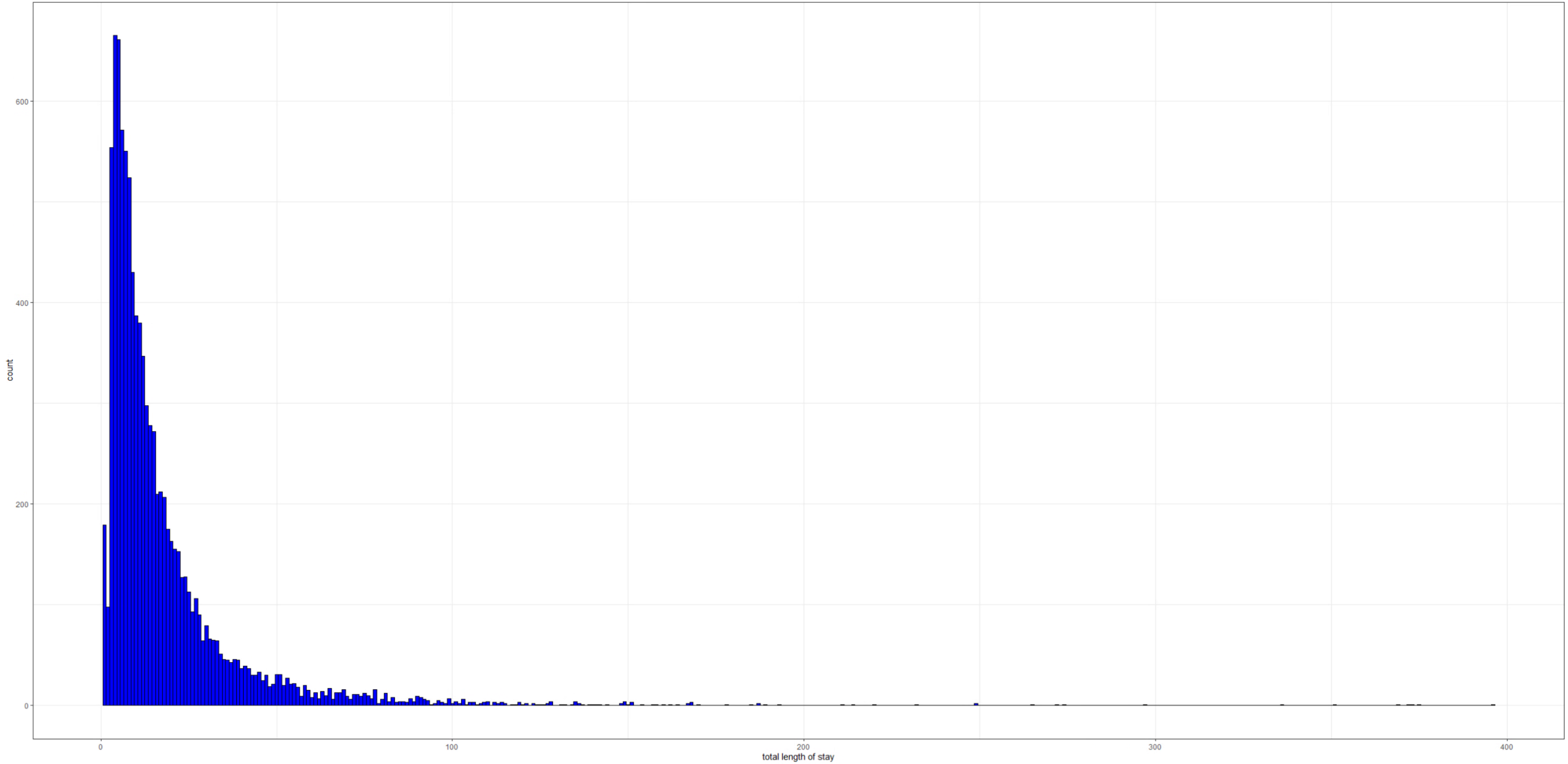
- Correct the time format
- Remove column with complete rate less than 10%
- Deal with Null value:replace them with the rest of the categories of corresponding column or simply remove the entire column
- Remove abnormal and meaningless values("Not recorded", "Not applicable")
- Convert some columns into dummy variables

Poor data quality& Impact on the forecasting accuracy

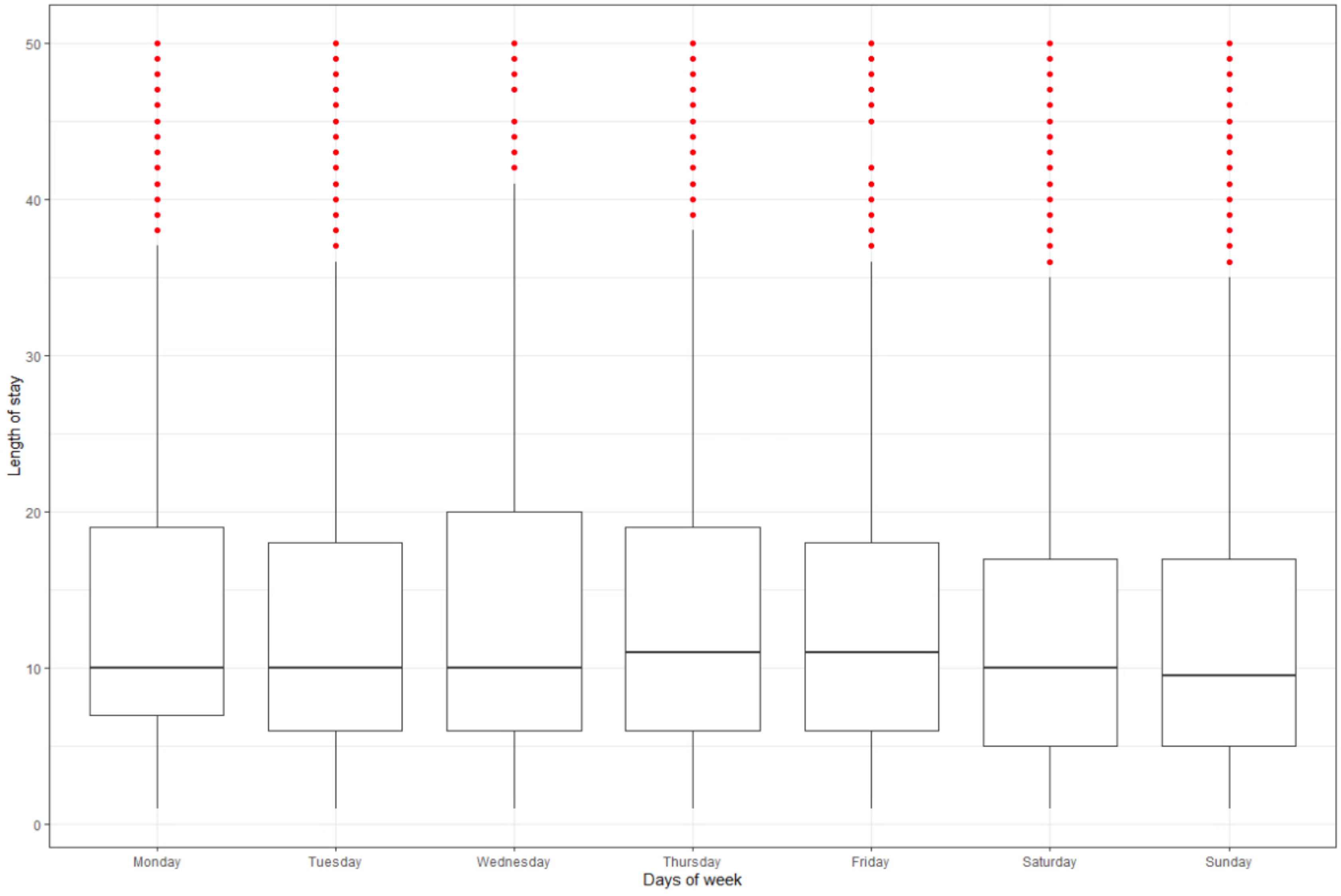
- 60% of the observation was removed
- Lack of some important time interval feature(time to CT scan, operation)

CONTENTS

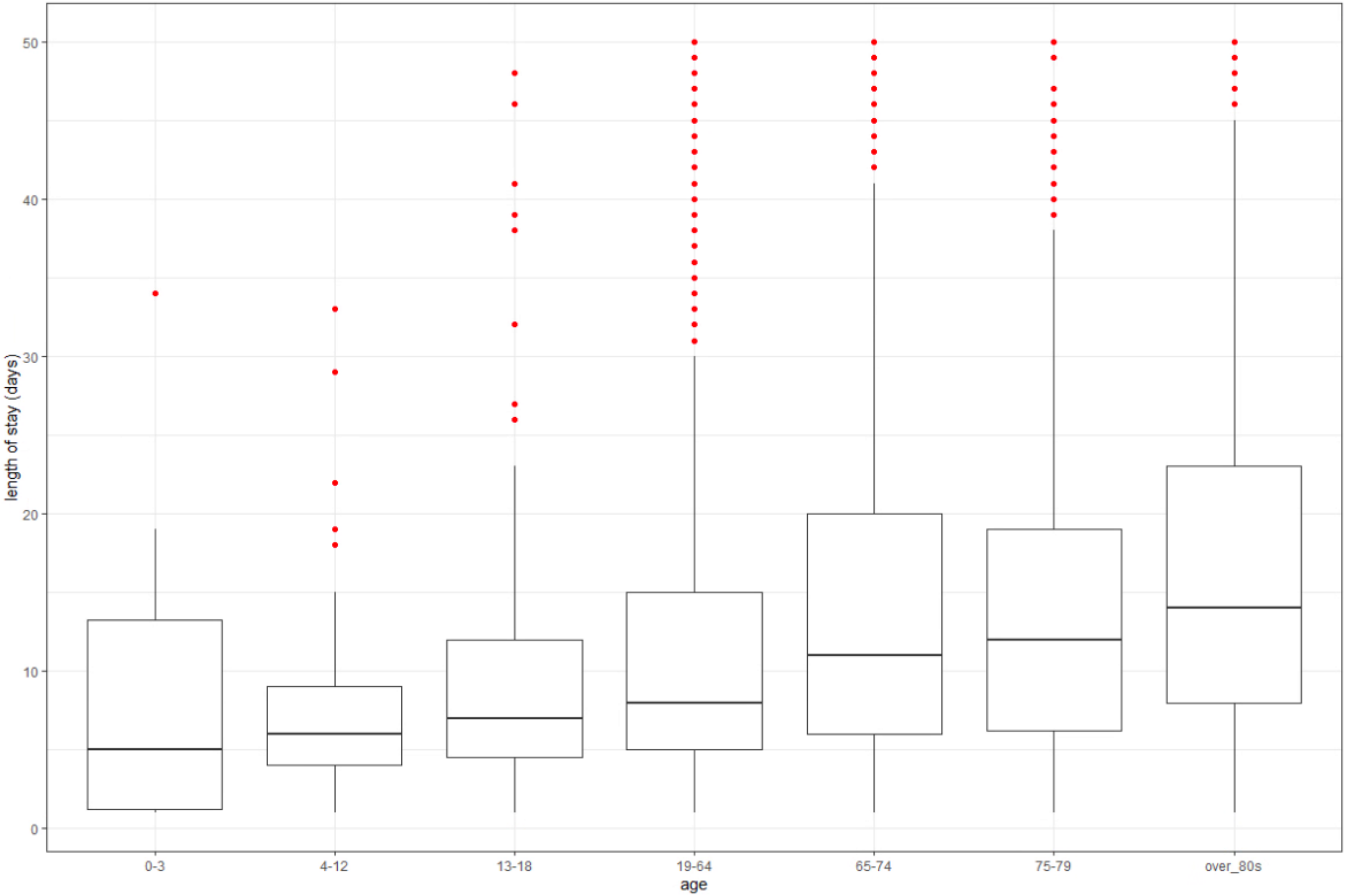
Data visualization: Distribution of the total length of stay



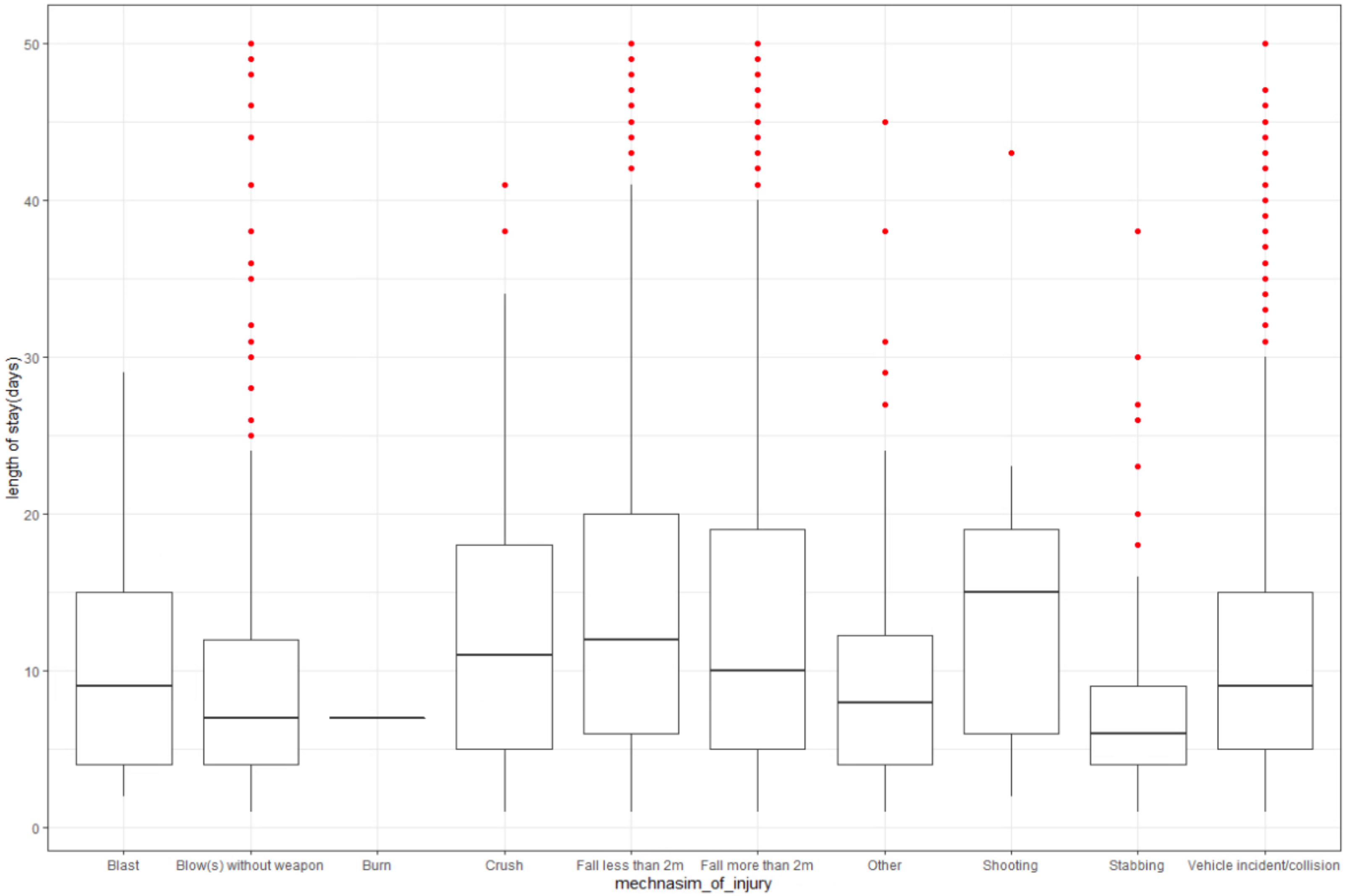
Data visualization:Covariance between LOS & patient arrival date



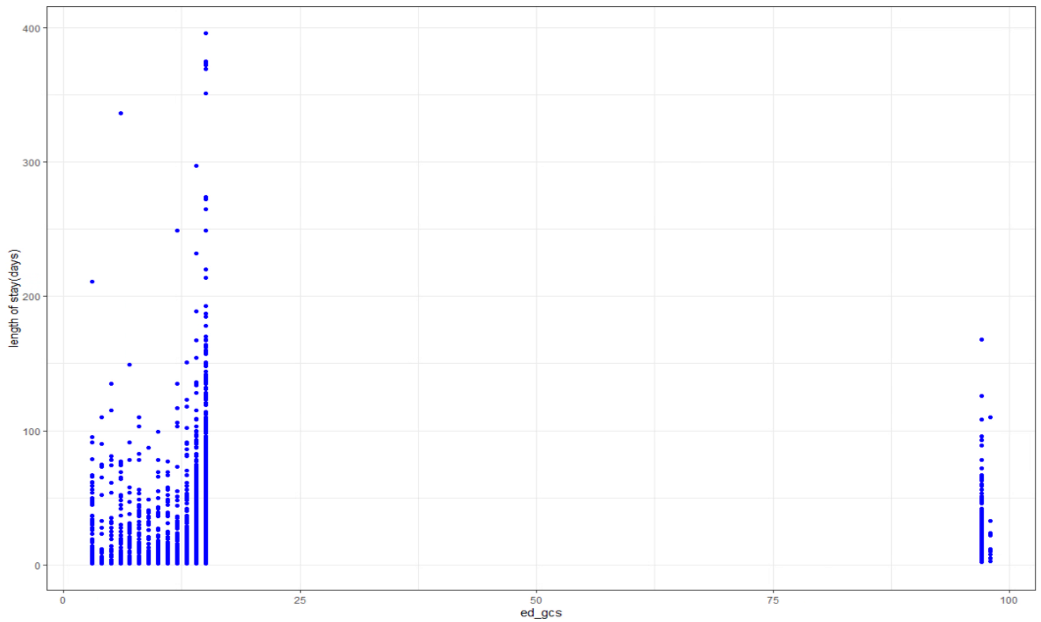
Data visualization: Covariance between LOS & age groups



Data visualization: Covariance between LOS & injury mechanism

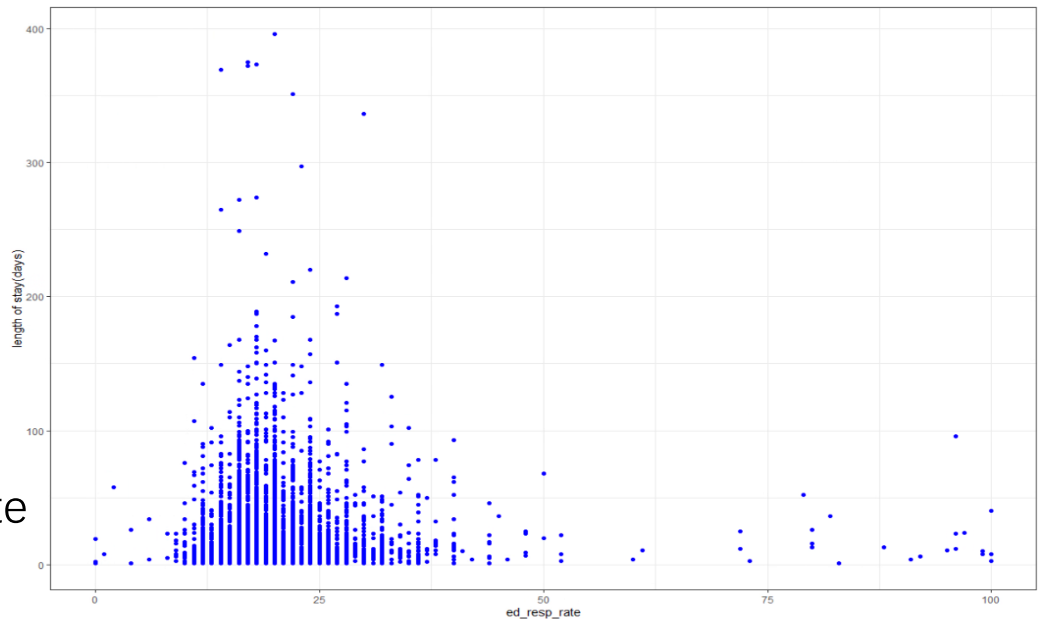


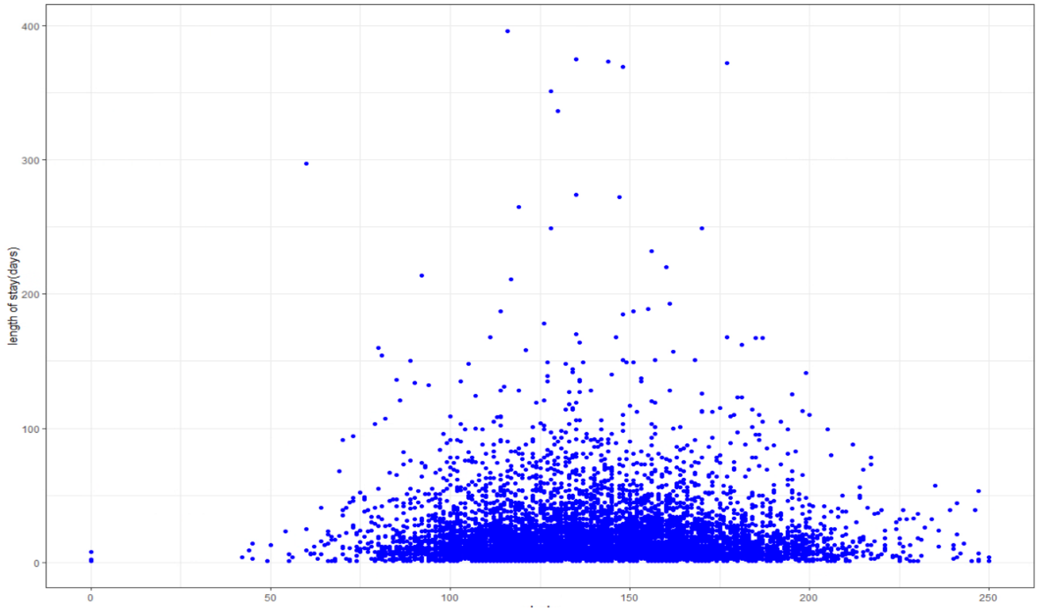
Data visualization: Covariance between LOS & numeric features



ed_gcs

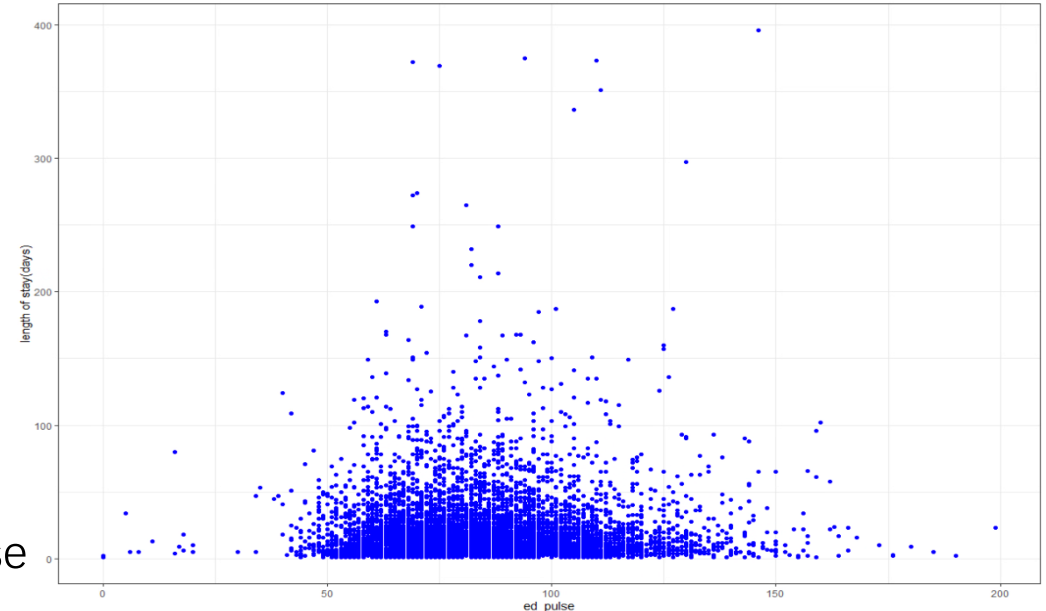
ed_resp_rate



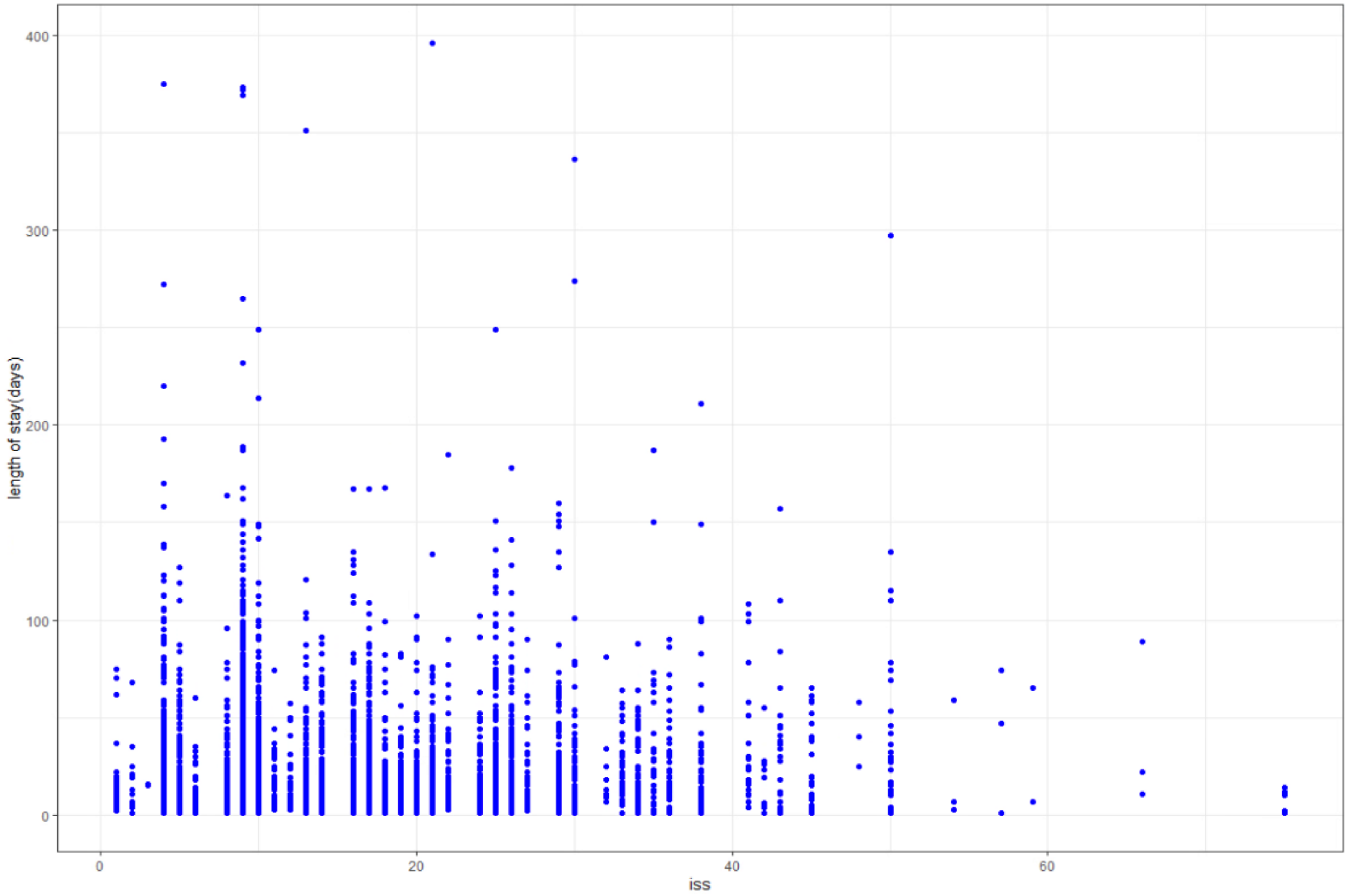


ed_sbp

ed_pulse



Data visualization: Covariation between LOS & Injury severity score (ISS)



CONTENTS

Methodology：

- **Lasso Regression**

$$\underbrace{\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2}_{RSS}+\lambda\underbrace{\sum_{j=1}^{p}|\beta_j|}_{\ell_1\ penalisation}$$

- **Ridge Regression**

$$\underbrace{\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2}_{RSS}+\underbrace{\lambda\sum_{j=1}^{p}\beta_j^2}_{\ell_2\ penalisation}$$

- **Elastic net regression**

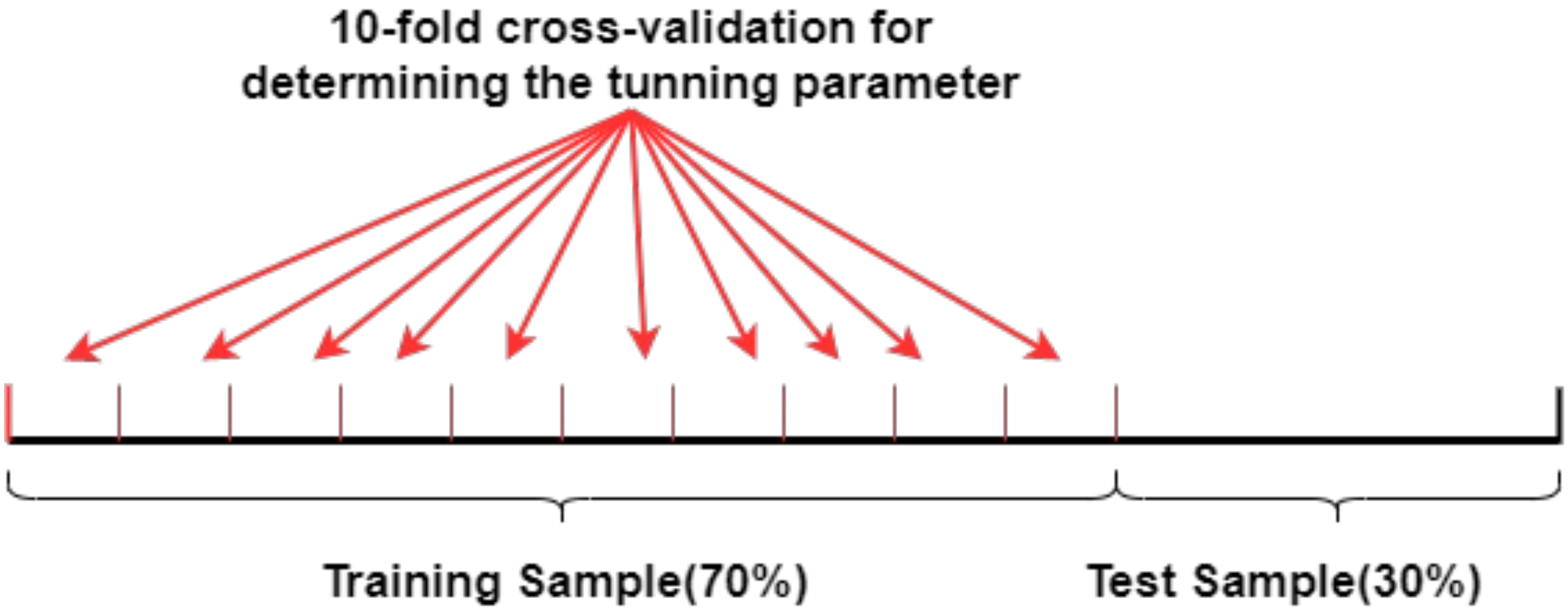
$$\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda1\sum_{j=1}^{p}\beta_j^2+\lambda2\sum_{j=1}^{p}|\beta_j|$$

- λ denotes the amount of shrinkage.

Reason for applying these regularization methods:

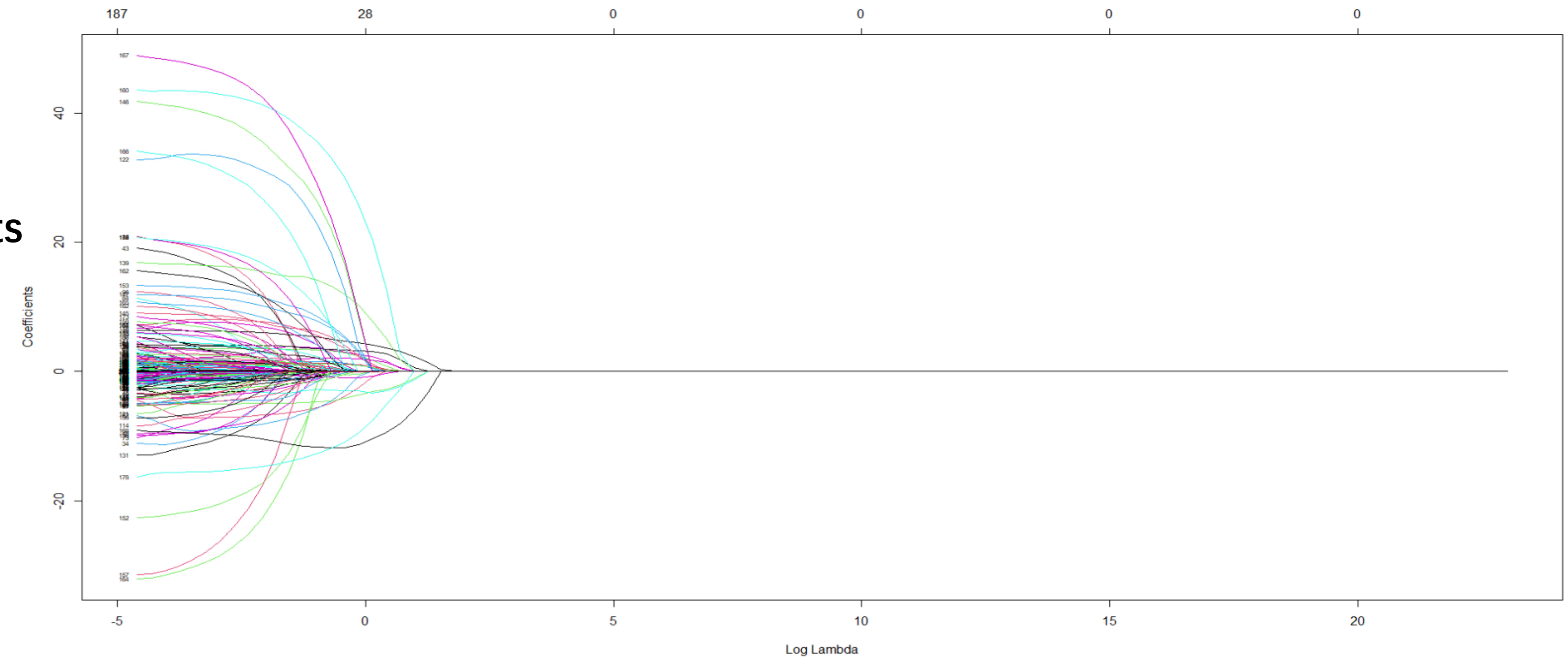
- Can be used to select important features from the dataset
- Shrinks the coefficients of less important features to exactly
- Independent features can be categorical, quantitative, or both

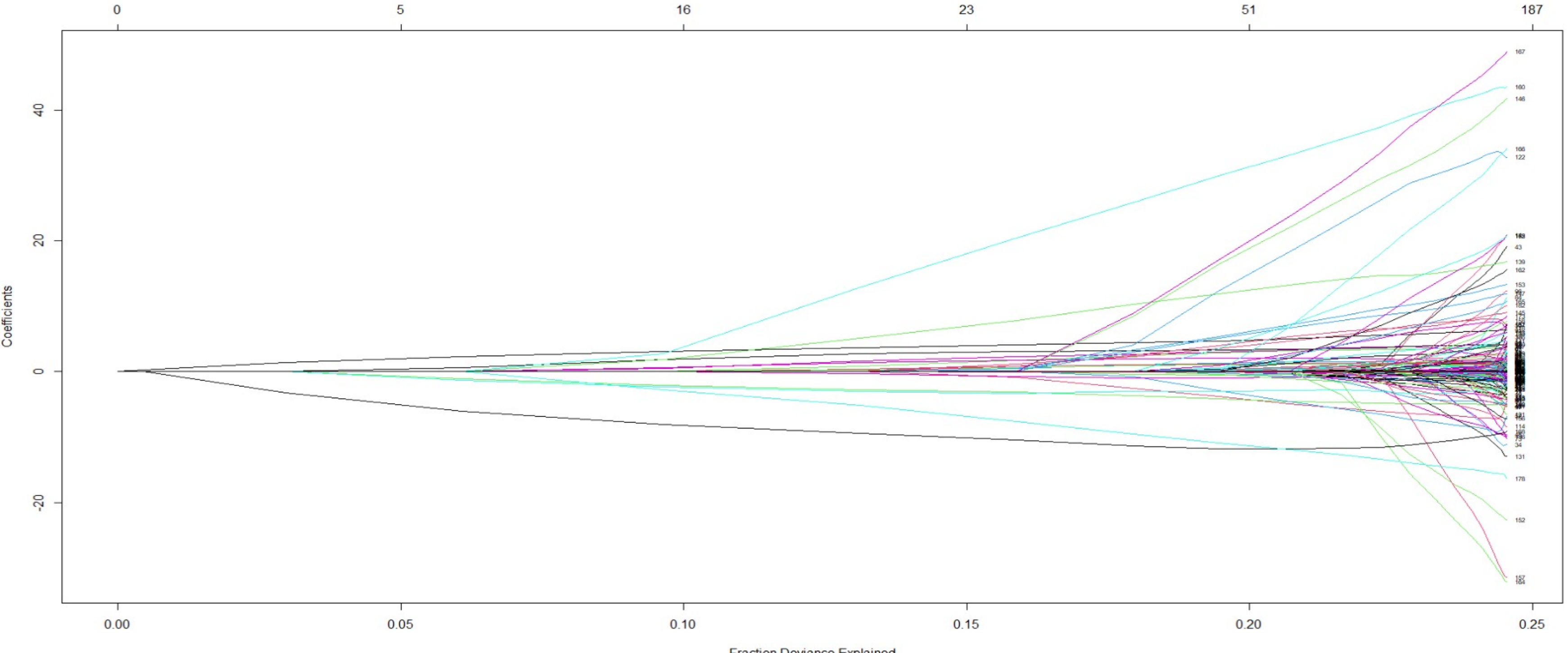
Methodology：technique for tunning parameter selection



Lasso model fit-in:

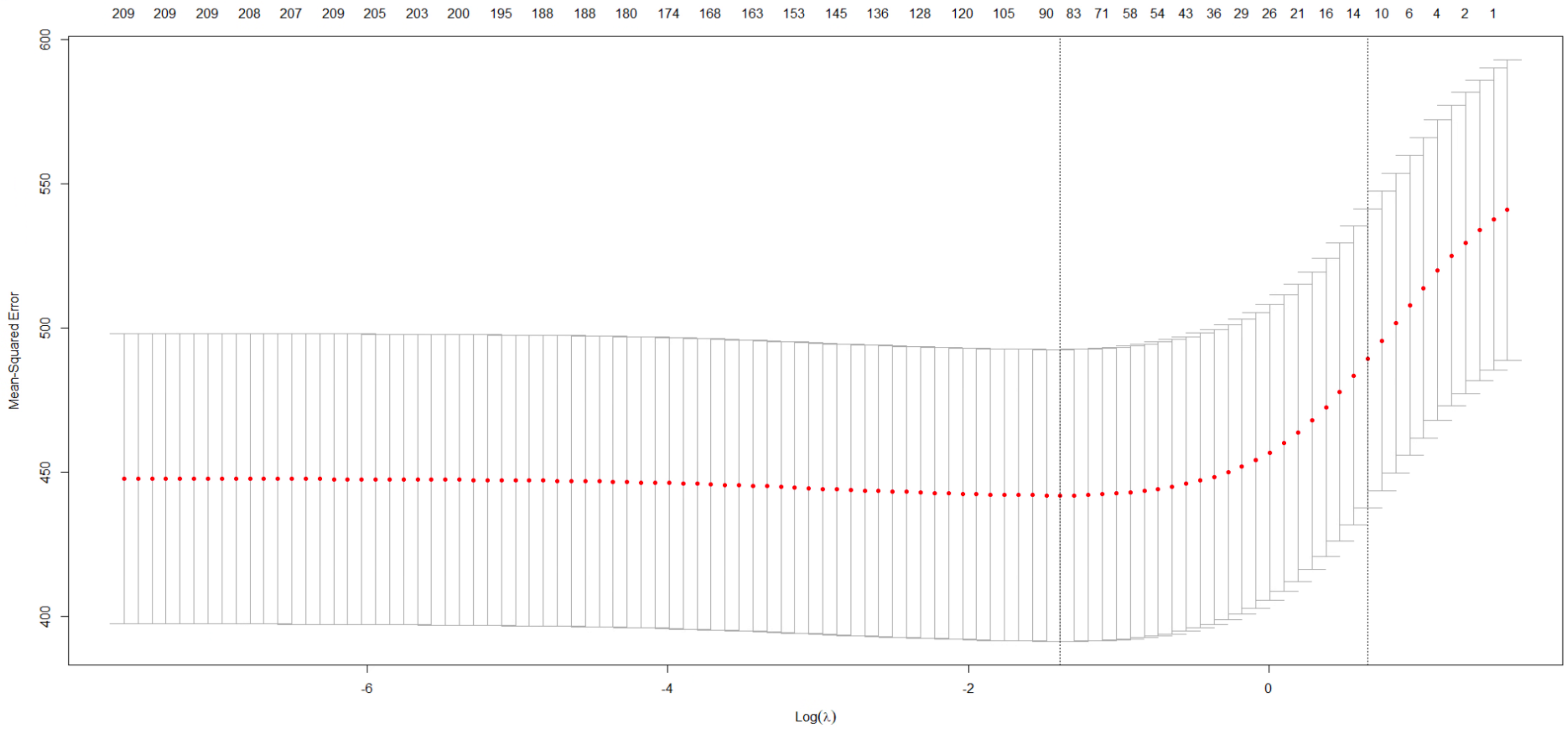
Lambda in relations to the coefficients





Goodness of fit

Lasso model fit-in: α=1， Choose the tuning parameter λ



The best λ value that minimizes the test MSE:0.2497.

Lasso model coefficient and forecasting accuracy

Intercept:14.3225

(Intercept)	14.3225
welshresident	0.3566
countryid	0.0000
casemtc	2.9966
mechnasim_of_injuryFall less than 2m	1.7443
injury_typePenetrating	-1.8064
locationHome	2.3702
locationInstitution	0.7224
locationOther Home (not patient's)	0.0592
arrival_modeHelicopter	-0.0193
highest_dgree_of_attendantParamedic	0.8745
age19-64	-4.2856
age75-79	1.2692
prealertYes	-0.1074
first_doctor_see_patientsConsultant	-0.2690
first_doctor_see_patients_datetimeTuesday	0.5274
total_ed_intubvent	1.1272
totalnumbers_of_operations	4.7972
abdomen	-0.0721
pelvis	1.3962
most_severeChest	-0.4963
most_severeLimbs	0.3518
ward1Coronary Care Unit (CCU)	-0.0440
ward1Geriatric	4.0309
ward1Level 4	25.4329
ward1Medical ward (inc. Pallative care)	0.2882
ward1Spinal injuries unit	6.4835
ward2Cardiothoracic	-1.4217
ward2General acute (inc. paediatric)	0.8541
ward2Level 2	-2.0063
ward2Major trauma ward	0.8589
ward2Medical ward (inc. Pallative care)	6.5495
ward2Neurosurgical ward	8.9831
ward2Orthopaedic (inc. paediatric)	1.8711
ward2Post Anaesthetic Care Unit (PACU)	-4.2429
ward2Surgical ward (inc. paediatric)	-0.5874
ward3Emergency Admissions Unit (EAU)	-1.9384

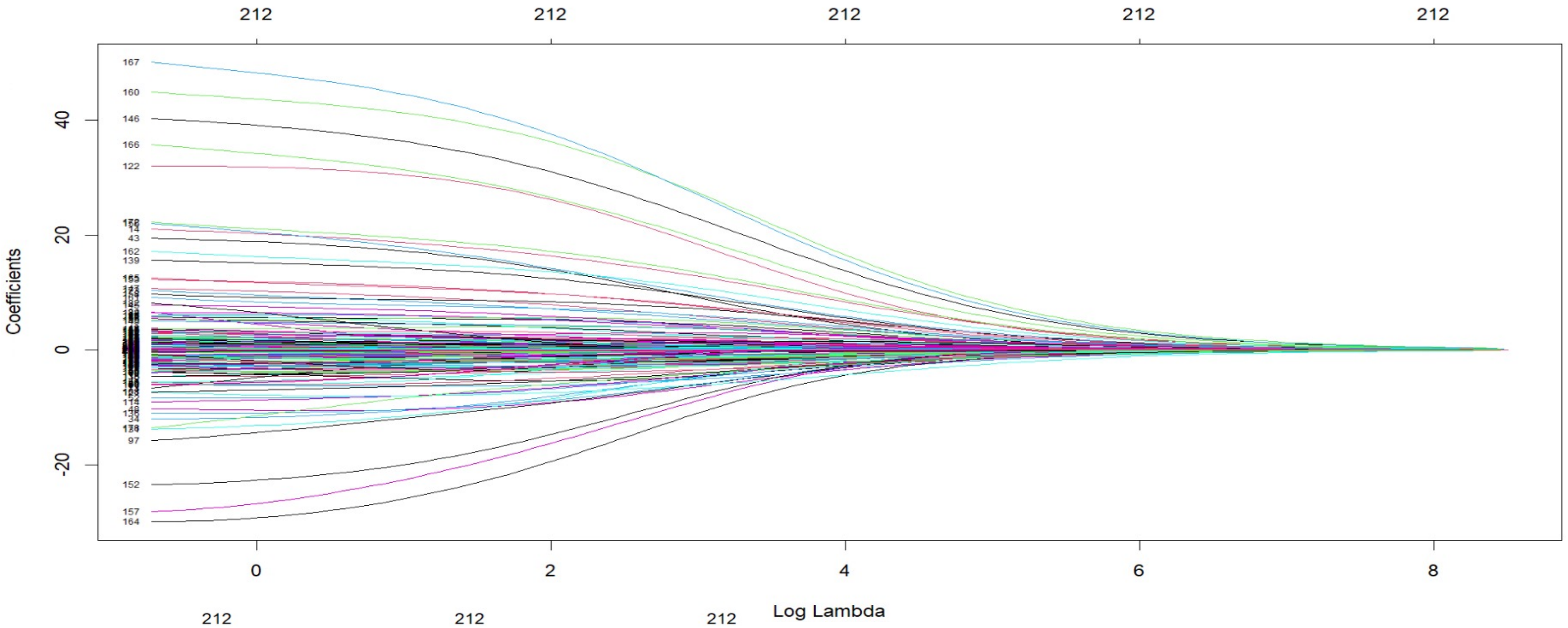
welshincident	0.4675
welshhospital	3.7861
mtc	0.8636
incident_dateSaturday	0.0815
mechnasim_of_injuryOther	0.0443
locationFarm	-0.8640
locationIndustrial	-1.2427
locationMountain	-1.5611
locationRoad	-0.6047
highest_dgree_of_attendantFY / Other	1.3218
age13-18	-6.7411
age4-12	-5.4383
ageover_80s	4.3928
ed_most_senior_doctor_datetimeSunday	-0.3096
first_doctor_see_patientsST 3+	0.1278
nice_headinjuty_cretriaYes	2.4809
numbers_of_operations1	-0.3409
type_of_transferTransfer Out	3.6575
spine	1.5742
limbs	0.0012
most_severeFace	-2.2338
ward1Cardiothoracic	-1.5079
ward1Emergency Admissions Unit (EAU)	-1.6907
ward1Level 3	2.5678
ward1Maxillofacial	-0.9144
ward1Orthopaedic (inc. paediatric)	2.2512
ward1Surgical ward (inc. paediatric)	-0.0265
ward2Emergency Admissions Unit (EAU)	-2.3682
ward2Geriatric	18.5301
ward2Level 3	2.8540
ward2Maxillofacial	-2.7834
ward2Neurosurgical rehabilitation ward	11.3555
ward2no_admission	-3.1175
ward2Plastic Surgery	-0.2932
ward2Spinal injuries unit	7.4505
ward3Cardiothoracic	0.2596
ward3General acute (inc. paediatric)	2.7558

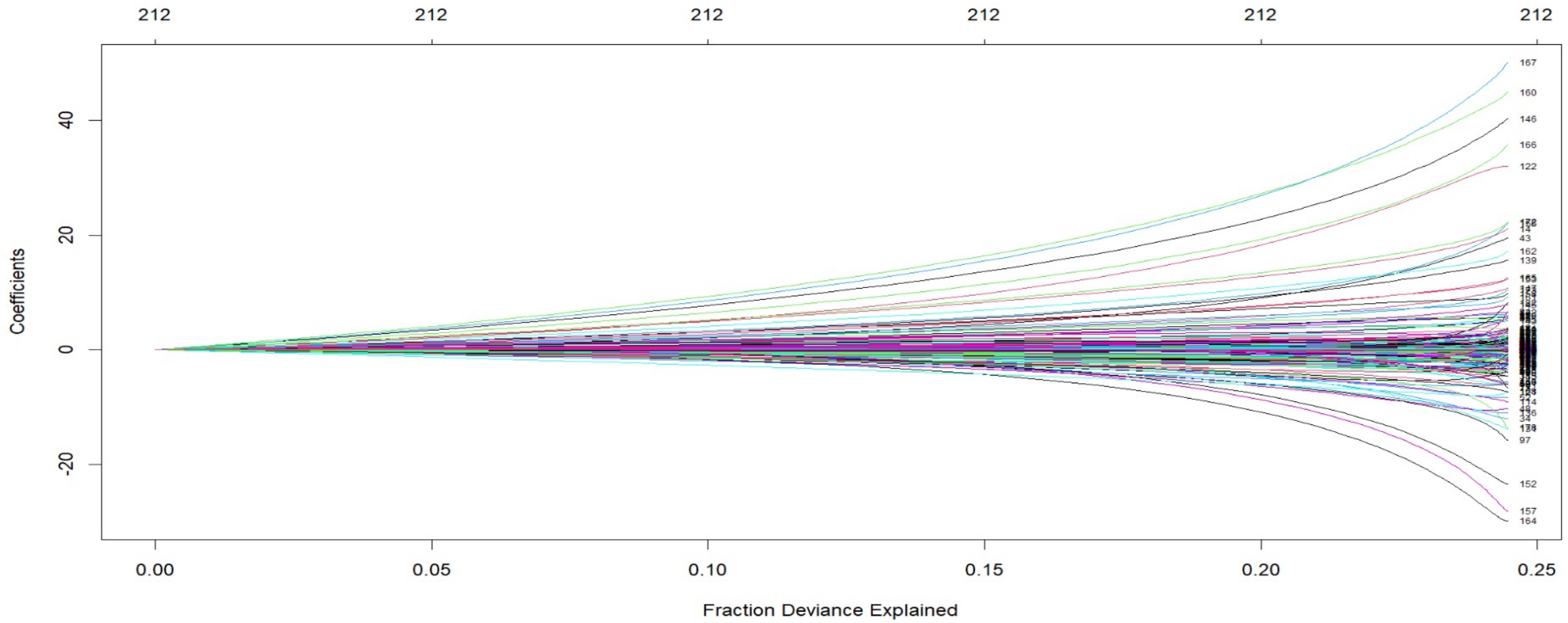
93 out of 214 variables with non-zero coefficients

Included features
welsh_resident
welsh_incident
welsh_hospital
casemtc
location
highest_degree_of_attendent
age
time for first_doctor_see_patients
type of transfer
ward type
most_sever injury
status of discharge
nice_head_injury_cretria
ed_pulse
ed_resp_rate
Ps14
case_known outcome
known outcome
prehospital_pulse
patient-arrival_time
txa_location

Ridge model fit-in:

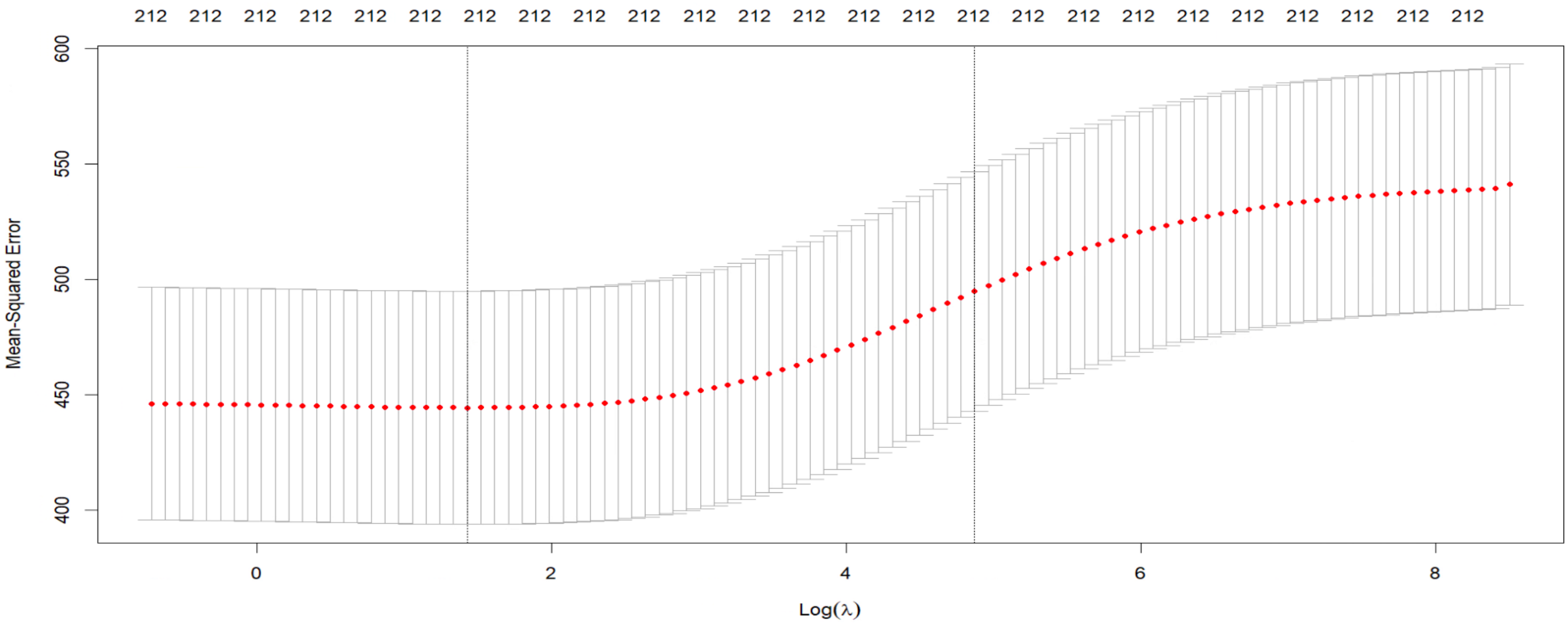
Lambda in relations to the coefficients





Goodness of fit

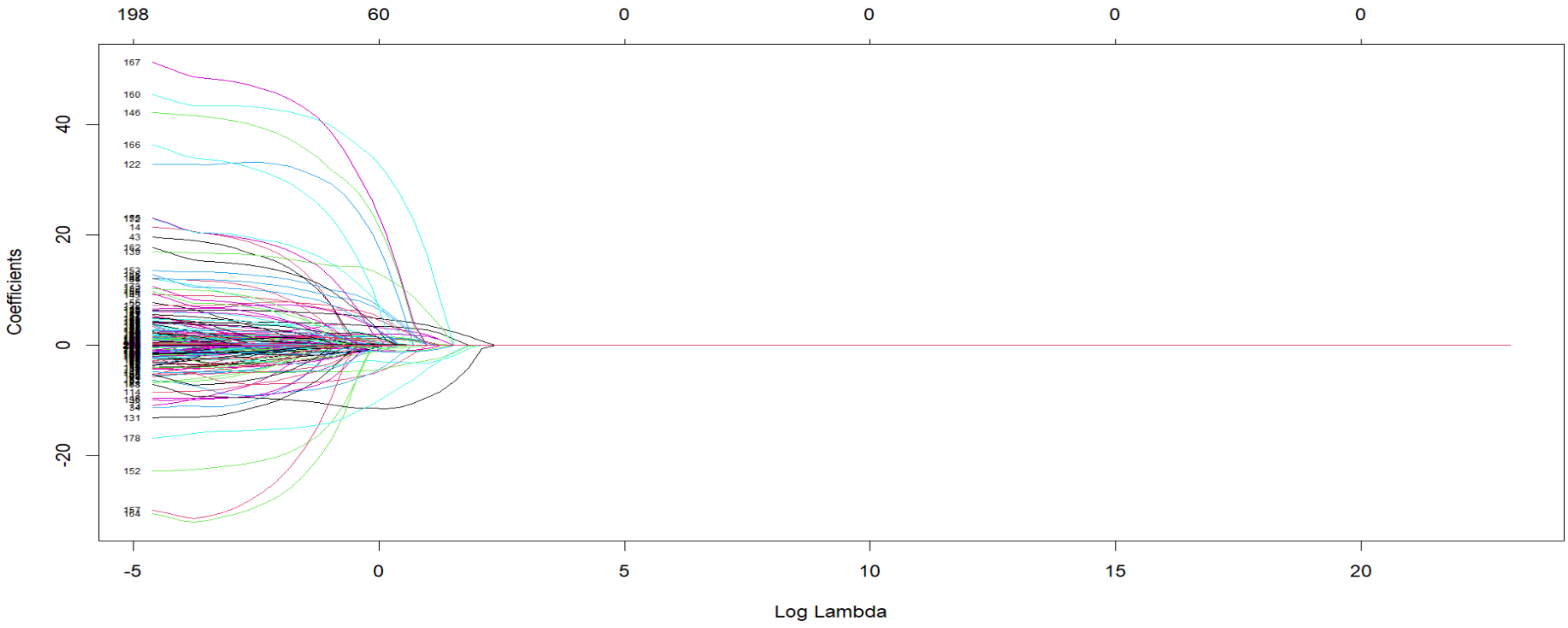
Ridge model fit-in: α=0， Choose the tuning parameter λ

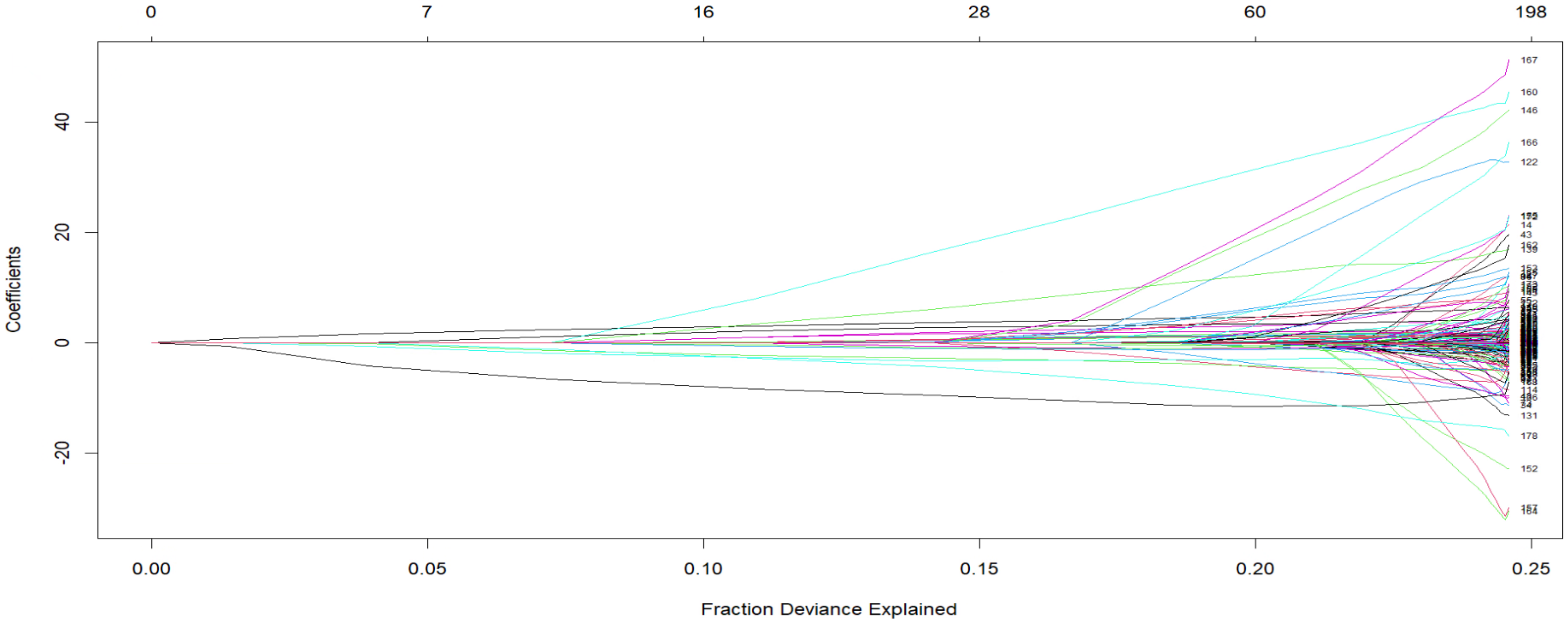


The best λ value that minimizes the test MSE:4.165675

Elastic-net model fit-in:

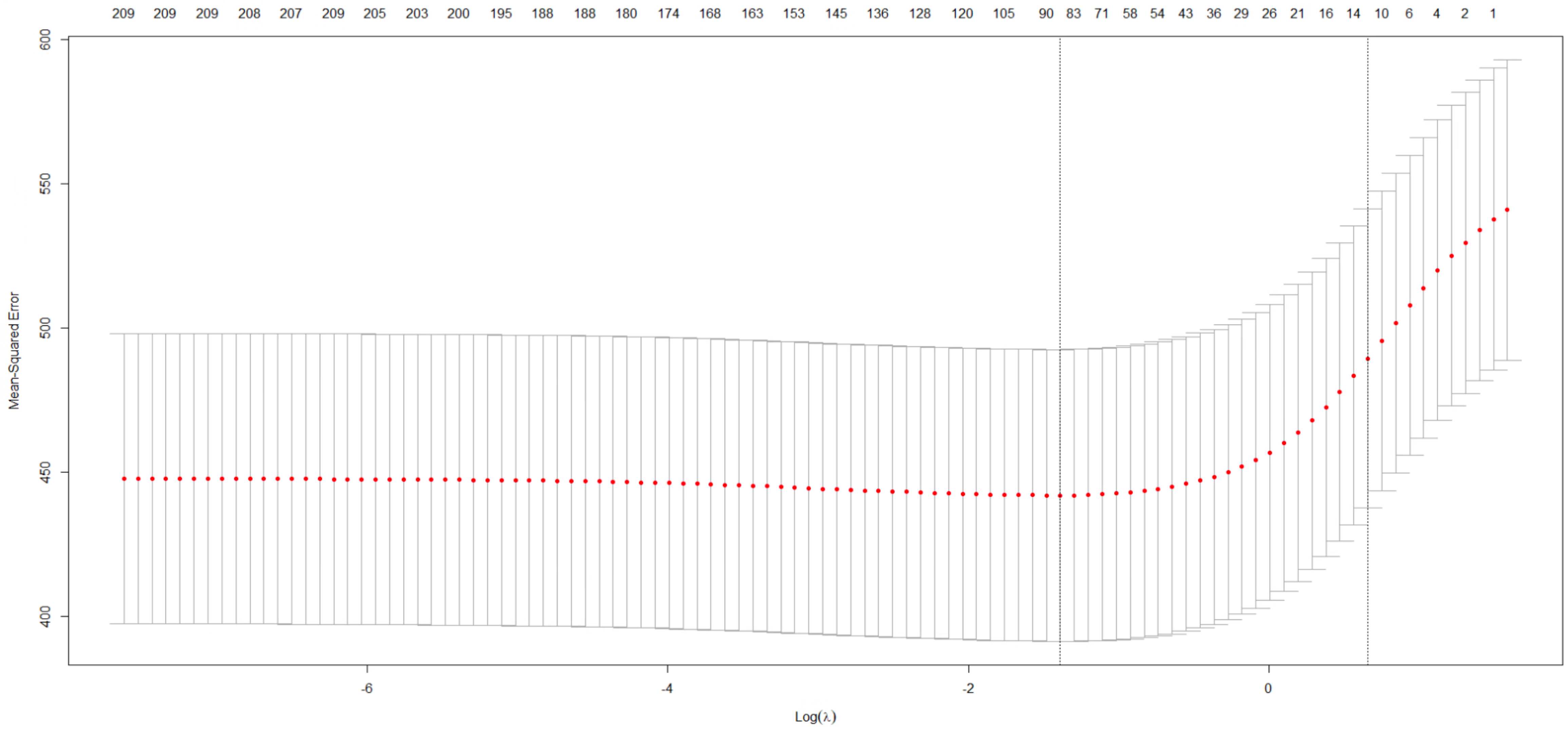
Lambda in relations to the coefficients





Goodness of fit

Elastic-net model fit-in: α=0.5,choose the tuning parameterλ



The best λ value that minimizes the test MSE:0.249725.

Elastic-net model coefficient and forecasting accuracy

Intercept:21.1424

(Intercept)	21.1424	welshincident	0.1302
welshresident	0.0060	welshhospital	3.4932
countryid	0.0000	casemtc	3.4191
mechnasim_of_injuryFall less than 2m	1.8320	mechnasim_of_injuryVehicle incident/collision	-0.1335
injury_typePenetrating	-1.1841	locationHome	2.6084
locationIndustrial	-0.0377	locationInstitution	0.0534
locationMountain	-0.0196	locationRoad	-0.0495
highest_dgree_of_attendantParamedic	0.4358	age13-18	-5.5660
age19-64	-4.0895	age4-12	-3.2629
age75-79	0.4906	ageover_80s	3.9858
ed_most_senior_doctor_datetimeSunday	-0.0144	first_doctor_see_patientsConsultant	-0.0455
first_doctor_see_patients_datetimeTuesday	0.0574	nice_headinjuty_cretriaYes	1.8266
total_ed_intubvent	0.8956	totalnumbers_of_operations	4.5578
type_of_transferTransfer Out	1.4091	spine	1.4369
pelvis	1.3191	most_severeChest	-0.7630
most_severeFace	-1.5028	ward1Cardiothoracic	-0.8656
ward1Emergency Admissions Unit (EAU)	-1.0023	ward1Geriatric	1.6497
ward1Level 3	1.8368	ward1Level 4	16.4339
ward1Maxillofacial	-0.0695	ward1Orthopaedic (inc. paediatric)	1.3335
ward1Spinal injuries unit	4.9603	ward1Surgical ward (inc. paediatric)	-0.0256
ward2Cardiothoracic	-0.1750	ward2Geriatric	17.3647
ward2Level 2	-0.2789	ward2Level 3	1.5933
ward2Maxillofacial	-0.4080	ward2Medical ward (inc. Pallative care)	5.3774
ward2Neurosurgical rehabilitation ward	5.0889	ward2Neurosurgical ward	7.5630
ward2no_admission	-2.8426	ward2Orthopaedic (inc. paediatric)	1.2598
ward2Spinal injuries unit	5.5021	ward2Surgical ward (inc. paediatric)	-0.0392
ward3Geriatric	28.9209	ward3Level 3	0.6762
ward3Medical ward (inc. Pallative care)	2.0958	ward3Neurosurgical rehabilitation ward	3.8134
ward3Neurosurgical ward	9.9236	ward3no_admission	-9.8171
ward3Spinal injuries unit	4.9850	iss	0.0993
ps14	-0.0553	status_of_dischargeDead	-0.3480
casedied	-11.9319	caseknownoutcome	0.3416
ed_pulse	0.0111	ed_resp_rate	0.0016

66 out of 214 variables with non-zero coefficients

Included features
welsh_resident
welsh_incident
welsh_hospital
casemtc
location
highest_degree_of_attendent
age
time for first_doctor_see_patients
type of transfer
ward type
most_sever injury
status of discharge
nice_head_injury_cretria
ed_pulse
ed_resp_rate
ps14

Revised-Elastic-net model: choose the tuning parameter α&λ

- Tune Length= 10 (α~(0.1-1))

alpha	lambda	RMSE	Rsquared	MAE
0.1	0.002272555	20.40848	0.19019952	11.51178
0.1	0.005249897	20.40828	0.19020938	11.51155
0.1	0.012127944	20.40360	0.19046074	11.50692
0.1	0.028017123	20.39525	0.19091230	11.49798
0.1	0.064723189	20.38304	0.19151337	11.48183
0.1	0.149518960	20.36339	0.19245845	11.45094
0.1	0.345408190	20.33077	0.19407586	11.39398
0.1	0.797937719	20.30316	0.19539403	11.31600
0.1	1.843339624	20.31361	0.19539691	11.26119
0.1	4.258353613	20.46778	0.19161455	11.35845
0.2	0.002272555	20.40900	0.19018402	11.51193
0.2	0.005249897	20.40544	0.19037503	11.50886
0.2	0.012127944	20.39907	0.19072218	11.50208
0.2	0.028017123	20.38822	0.19128934	11.48900
0.2	0.064723189	20.37177	0.19209747	11.46385
0.2	0.149518960	20.34106	0.19369168	11.41593
0.2	0.345408190	20.30620	0.19546250	11.33782
0.2	0.797937719	20.29808	0.19589803	11.26577
0.2	1.843339624	20.38865	0.19248992	11.29304
0.2	4.258353613	20.76912	0.17923987	11.63437
0.3	0.002272555	20.40805	0.19023228	11.51101
⋮				
1.0	0.002272555	20.40205	0.19055932	11.50474
1.0	0.005249897	20.39221	0.19110184	11.49369
1.0	0.012127944	20.37741	0.19185900	11.47248
1.0	0.028017123	20.34850	0.19340751	11.43016
1.0	0.064723189	20.30983	0.19551112	11.35538
1.0	0.149518960	20.28919	0.19654588	11.27234
1.0	0.345408190	20.35378	0.19293487	11.26665
1.0	0.797937719	20.63470	0.17980237	11.48895
1.0	1.843339624	21.51360	0.12552102	12.35212
1.0	4.258353613	22.44415	0.05091961	13.26945

- **RMSE** was used to select the optimal model using the smallest value

The final used for the model was α=1 and $\lambda = 0.149519$

Revised Elastic-net model coefficient and forecasting accuracy

Intercept:7.038

(Intercept)	7.038
welshresident	0.555
countryid	0.233
casemtc	1.920
mechnasim_of_injuryFall less than 2m	1.627
injury_typePenetrating	-1.903
locationHome	2.276
locationInstitution	1.226
locationOther	0.360
locationRoad	-0.746
arrival_modeHelicopter	-0.073
highest_dgree_of_attendantFY / Other	7.761
highest_dgree_of_attendantParamedic	0.956
age13-18	-7.102
age4-12	-6.460
ageover_80s	4.584
ed_most_senior_doctorOther	0.289
ed_most_senior_doctor_datetimeSunday	-0.426
first_doctor_see_patientsConsultant	-0.320
first_doctor_see_patients_datetimeTuesday	0.923
gcs	-0.019
numbers_of_operations1	-1.213
type_of_transferTransfer Out	7.921
abdomen	-0.146
pelvis	1.398
most_severeChest	-0.367
most_severeLimbs	0.711
ward1Coronary Care Unit (CCU)	-1.416
ward1Geriatric	5.257
ward1Level 4	29.234
ward1Medical ward (inc. Pallative care)	0.881
ward1Orthopaedic (inc. paediatric)	2.711
ward2Cardiothoracic	-1.590
ward2General acute (inc. paediatric)	2.428
ward2Level 2	-2.651
ward2Level 3s	4.616
ward2Maxillofacial	-3.242

welshincident	0.694
welshhospital	3.632
mtc	2.256
incident_dateSaturday	0.457
mechnasim_of_injuryOther	0.809
locationFarm	-1.462
locationIndustrial	-1.969
locationMountain	-2.571
locationOther Home (not patient's)	0.448
locationWater	-0.273
arrival_modeOther	-1.378
highest_dgree_of_attendantNo grade recorded	-0.471
highest_dgree_of_attendantST, 4+	-1.738
age19-64	-4.303
age75-79	1.687
prealertYes	-0.280
ed_most_senior_doctorST year unknown	0.006
ed_most_senior_doctor_datetimeWednesday	0.118
first_doctor_see_patientsST 3+	0.350
nice_headinjuty_cretriaYes	2.868
total_ed_intubvent	1.334
totalnumbers_of_operations	5.076
face	-0.111
spine	1.675
limbs	0.105
most_severeFace	-2.305
ward1Cardiothoracic	-1.914
ward1Emergency Admissions Unit (EAU)	-1.991
ward1Level 3	2.825
ward1Maxillofacial	-0.883
ward1no_admission	-0.084
ward1Spinal injuries unit	7.226
ward2Emergency Admissions Unit (EAU)	-5.556
ward2Geriatric	19.503
ward2Level 3	4.001
ward2Major trauma ward	2.895
ward2Medical ward (inc. Pallative care)	7.576

Included features
welsh_resident
welsh_incident
welsh_hospital
casemtc
location
highest_degree_of_attendent
age
time for first_doctor_see_patients
type of transfer
ward type
most_sever injury
status of discharge
nice_head_injury_cretria
ed_pulse
ed_resp_rate
Ps14
case_known outcome
known outcome
prehospital_pulse
patient-arrival_time
txa_location
AIS maximum severity in Abdomen
gcs
arrival_mode
prealert

110 out of 214 variables with non-zero coefficients

Evaluation of the model performance

Model	Rsquared	RMSE_values
Linear Regression	0.1686	21.1628
Ridge Regression	0.1857	20.9433
Lasso Regression	0.1817	20.9957
Elastic Regression	0.1825	20.9846
Elastic Regression_r1	0.1965	21.8642

CONTENTS

Summary of the model

01 Loss of some important time interval feature

02 Plenty of outliers brings high difficulty in fit-in the forecasting model

03 Poor performance of the set of regularization forecasting model

Future research perspective

Acquire the latest dataset with complete time interval features for the modification of the forecasting model

Explore other forecasting methodologies (GAM, deep learning)

A further literature review for identify length of stay forecasting predictors and methodologies

Consider the impact of capacity variable (utilization of the medical resources) to the length of stay forecasting

Thank you for your listening

ANY QUESTIONS