

UNDERSTANDING THE DATASET

DATE	24 October 2022
Team ID	PNT2022TMID02935
Project	Project - Analytics For Hospitals' Health-Care Data

The collected dataset is divided into training and testing dataset for training and testing the dataset.

training data ☆ 📄 🌐

File Edit View Insert Format Data Tools Extensions Help *Last edit was 3 days ago*

100% \$ % .00 123 ▾ Default (A1) 10 B I S A 🗑️ 📊 📈 📉 📊 📈 📉 📊 📈 📉 📊 📈 📉

A1	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	case_id	Hospital_code	Hospital_type_o	City_Code_Hos	Hospital_region	Available Extra	Department	Ward_Type	Ward_Facility_C	Bed Grade	patientid	City_Code_Patie	Type of Admissi	Severity of Illnes
1	case_id													
2	1	8 c		3 Z			3 radiotherapy	R	F		2	31397	7 Emergency	Extreme
3	2	2 c		5 Z			2 radiotherapy	S	F		2	31397	7 Trauma	Extreme
4	3	10 e		1 X			2 anesthesia	S	E		2	31397	7 Trauma	Extreme
5	4	26 b		2 Y			2 radiotherapy	R	D		2	31397	7 Trauma	Extreme
6	5	26 b		2 Y			2 radiotherapy	S	D		2	31397	7 Trauma	Extreme
7	6	23 a		6 X			2 anesthesia	S	F		2	31397	7 Trauma	Extreme
8	7	32 f		9 Y			1 radiotherapy	S	B		3	31397	7 Emergency	Extreme
9	8	23 a		6 X			4 radiotherapy	Q	F		3	31397	7 Trauma	Extreme
10	9	1 d		10 Y			2 gynecology	R	B		4	31397	7 Trauma	Extreme
11	10	10 e		1 X			2 gynecology	S	E		3	31397	7 Trauma	Extreme
12	11	22 g		9 Y			2 radiotherapy	S	B		2	31397	7 Urgent	Extreme
13	12	26 b		2 Y			4 radiotherapy	R	D		1	31397	7 Urgent	Extreme
14	13	16 c		3 Z			2 radiotherapy	R	A		3	31397	7 Emergency	Extreme
15	14	9 d		5 Z			3 radiotherapy	S	F		3	31397	7 Urgent	Extreme
16	15	6 a		6 X			4 gynecology	Q	F		3	63418	8 Emergency	Extreme
17	16	6 a		6 X			3 gynecology	Q	F		3	63418	8 Emergency	Extreme
18	17	23 a		6 X			4 radiotherapy	Q	F		3	63418	8 Urgent	Extreme
19	18	29 a		4 X			4 anesthesia	S	F		3	63418	8 Emergency	Extreme
20	19	32 f		9 Y			4 radiotherapy	S	B		2	63418	8 Trauma	Extreme
21	20	12 a		9 Y			4 radiotherapy	Q	B		2	63418	8 Trauma	Extreme
22	21	16 c		3 Z			2 anesthesia	S	A		3	63418	8 Trauma	Extreme

training data

File Edit View Insert Format Data Tools Extensions Help [Last edit was 3 days ago](#)

A1	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
	Department	Ward_Type	Ward_Facility_C	Bed Grade	patientid	City_Code_Patie	Type of Admissio	Severity of Illnes	Visitors with Pati	Age	Admission_Dept	Stay		
2	radiotherapy	R	F		2	31397	7 Emergency	Extreme	2	51-60	4911	0-10		
3	radiotherapy	S	F		2	31397	7 Trauma	Extreme	2	51-60	5954	41-50		
4	anesthesia	S	E		2	31397	7 Trauma	Extreme	2	51-60	4745	31-40		
5	radiotherapy	R	D		2	31397	7 Trauma	Extreme	2	51-60	7272	41-50		
6	radiotherapy	S	D		2	31397	7 Trauma	Extreme	2	51-60	5558	41-50		
7	anesthesia	S	F		2	31397	7 Trauma	Extreme	2	51-60	4449	11-20		
8	radiotherapy	S	B		3	31397	7 Emergency	Extreme	2	51-60	6167	0-10		
9	radiotherapy	Q	F		3	31397	7 Trauma	Extreme	2	51-60	5571	41-50		
10	gynecology	R	B		4	31397	7 Trauma	Extreme	2	51-60	7223	51-60		
11	gynecology	S	E		3	31397	7 Trauma	Extreme	2	51-60	6056	31-40		
12	radiotherapy	S	B		2	31397	7 Urgent	Extreme	2	51-60	5797	21-30		
13	radiotherapy	R	D		1	31397	7 Urgent	Extreme	2	51-60	5993	11-20		
14	radiotherapy	R	A		3	31397	7 Emergency	Extreme	2	51-60	5141	0-10		
15	radiotherapy	S	F		3	31397	7 Urgent	Extreme	2	51-60	8477	21-30		
16	gynecology	Q	F		3	63418	8 Emergency	Extreme	2	71-80	2685	0-10		
17	gynecology	Q	F		3	63418	8 Emergency	Extreme	2	71-80	9398	0-10		
18	radiotherapy	Q	F		3	63418	8 Urgent	Extreme	4	71-80	2933	0-10		
19	anesthesia	S	F		3	63418	8 Emergency	Extreme	2	71-80	5342	11-20		
20	radiotherapy	S	B		2	63418	8 Trauma	Extreme	2	71-80	7442	21-30		
21	radiotherapy	Q	B		2	63418	8 Trauma	Extreme	2	71-80	5155	31-40		
22	anesthesia	S	A		3	63418	8 Trauma	Extreme	2	71-80	8181	31-40		

Columns in dataset (Parameters):

The collected data is details of the patients entered at the time of admission with the length of stay of the patients. The dataset has various important parameters which can be used to predict the Length Of Stay (LOS) of patients. Below is the description of each column or parameter.

Column name (Parameter)	Description
Case ID (case_id)	It is identity number given by hospital administration for the case of the patient during the time of admission
Hospital Code (Hospital_code)	It is the code (identity number) given to the hospital to uniquely identify the hospital.
Hospital type code (Hospital_type_code)	It is the unique code given to the type of hospital. This tells the hospital type.
Hospital city code (city_code_hospital)	It is the code given to the city where the hospital is located
Hospital region code (Hospital_region_code)	It is the code given to the region where the hospital is located.
Rooms available (Available_extra_room)	It will display the number of rooms that are still unoccupied, i.e, rooms available.
Department (Department)	The department that is overlooking the patients. The department to which the patient is admitted. For example radiotherapy, anesthesia, gynecology, TB & Chest disease, surgery.
Ward type (Ward_Type)	The unique code given to the type of ward to which the patient admitted

Ward facility code (Ward_Facility_Code)	The unique code given to the facility in the ward to which the patient is admitted.
Condition of the Bed (Bed_Grade)	It is the quality or condition of the bed in the ward to which the patient is admitted.
Patient ID (patient_id)	It is the unique identity value given to the patient at the time of admission
Patients' city code (City_Code_Patient)	It is the unique identity code given to the city from where the patient is from.
Type of admission	It is the admission type registered in the hospital at the time of admission.
Severity of Illness	It is the severity level of the patients' illness at the time of admission
Visitor with the patients	Number of the visitors with the patients to take extra care over them.
Age	It is the age of patients. It is given in period 0-10,11-20,21-20 to 91-100
Admission deposit amount (Admission_Deposit)	It is the deposit amount that the patient paid at the time of admission. The deposit amount will depend on various above features or parameters like department, ward type, Bed Grade, Type of Admission, Ward facility, severity of illness, hospital code.
Stay	It is the Length Of Stay (LOS) of patients. It is given as a period 0-10,11-20 to 91-100 Or More than hundred.

Total number of record in the dataset:

The dataset has

Parameter or Columns : 18

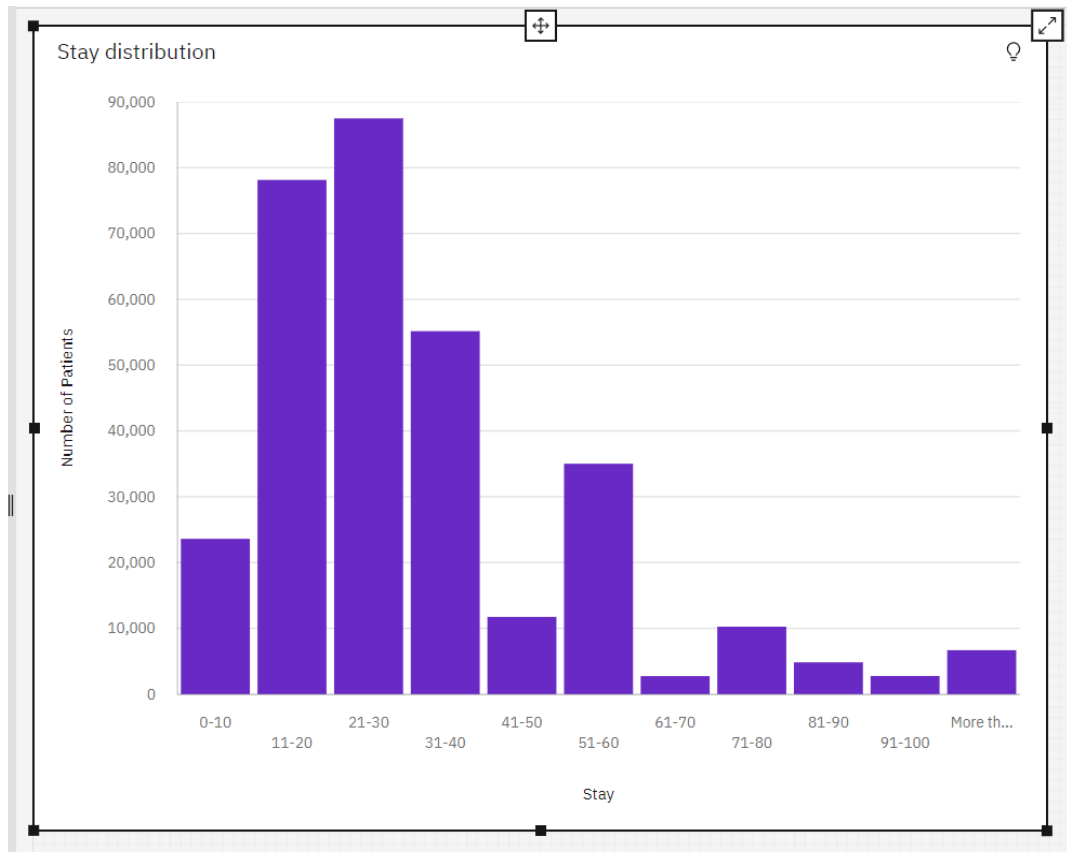
Rows or Records : 318438

Unique values of each column or parameter:

Column name	Unique values
Case id	from 1 to 318438
Hospital_code	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31
Hospital_type_code	a, b, c, d, e, f, g
City_Code_Hospital	1, 2, 3, 4, 5, 6, 7, 9, 19, 11, 13
Hospital_region_code	X, Y, Z
Available Extra Rooms	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 20, 21, 24
Unique Values for Department	Radiotherapy, anesthesia, gynecology, TB & Chest disease, surgery
Ward_Type	P, Q, R, S, T, U
Ward_Facility_Code	A, B, C, D, E, F
Bed Grade	0, 1, 2, 3, 4
patientid	31397, 63418, 8088 ... 125235, 91081, 21641

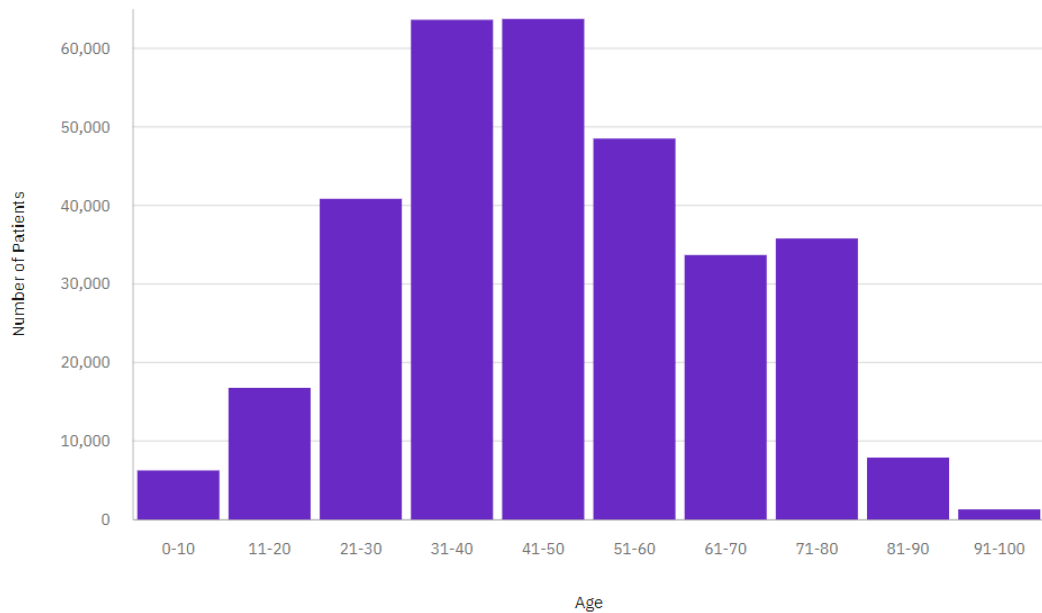
City_Code_Patient	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38
Type of Admission	Emergency, Trauma, Urgent
Severity of Illness	Extreme, Moderate, Minor
Visitors	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 32
Age	0-10, 11-20, 21-20, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, 90-100, 100 or more
Admission_Deposit	4911, 5954, 4745, ... ,1937, 9439, 2349
Values for Stay	0-10, 11-20, 21-20, 31-40, 41-50, 51-60, 61-70, 71-80, 81-90, 90-100, more than 100 days

Stay value distribution:

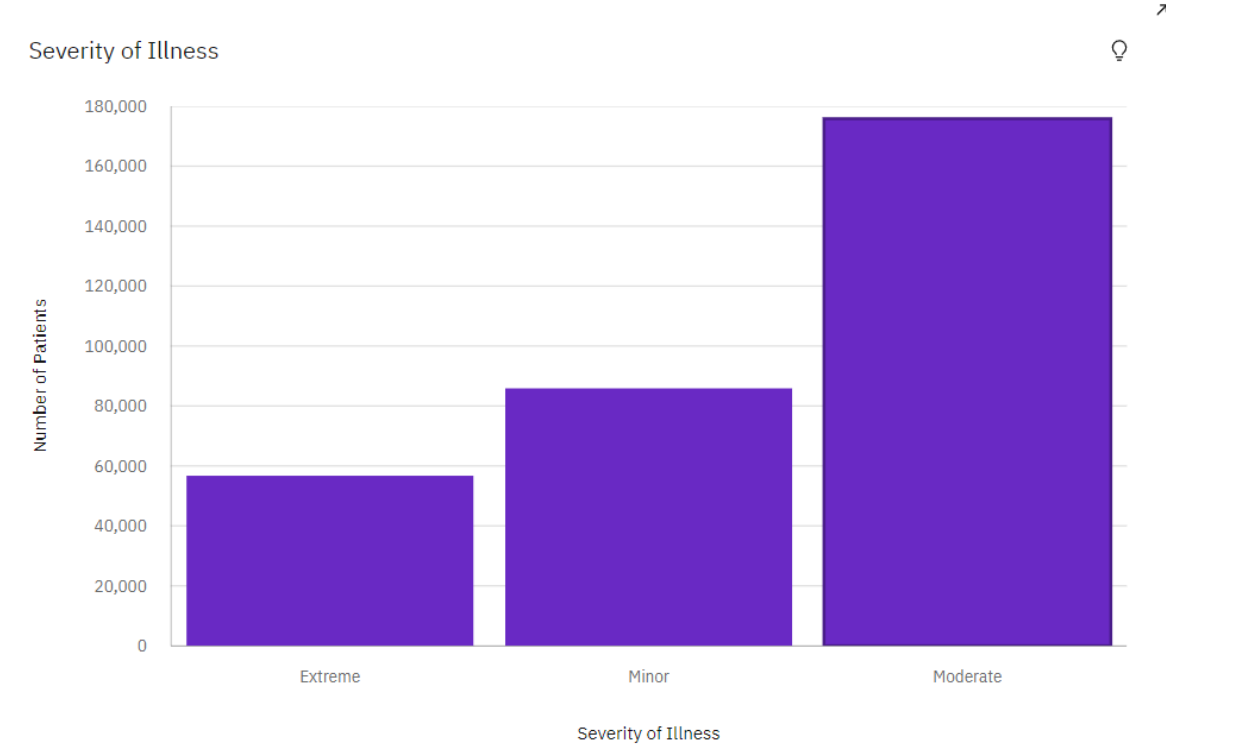


Age value distribution:

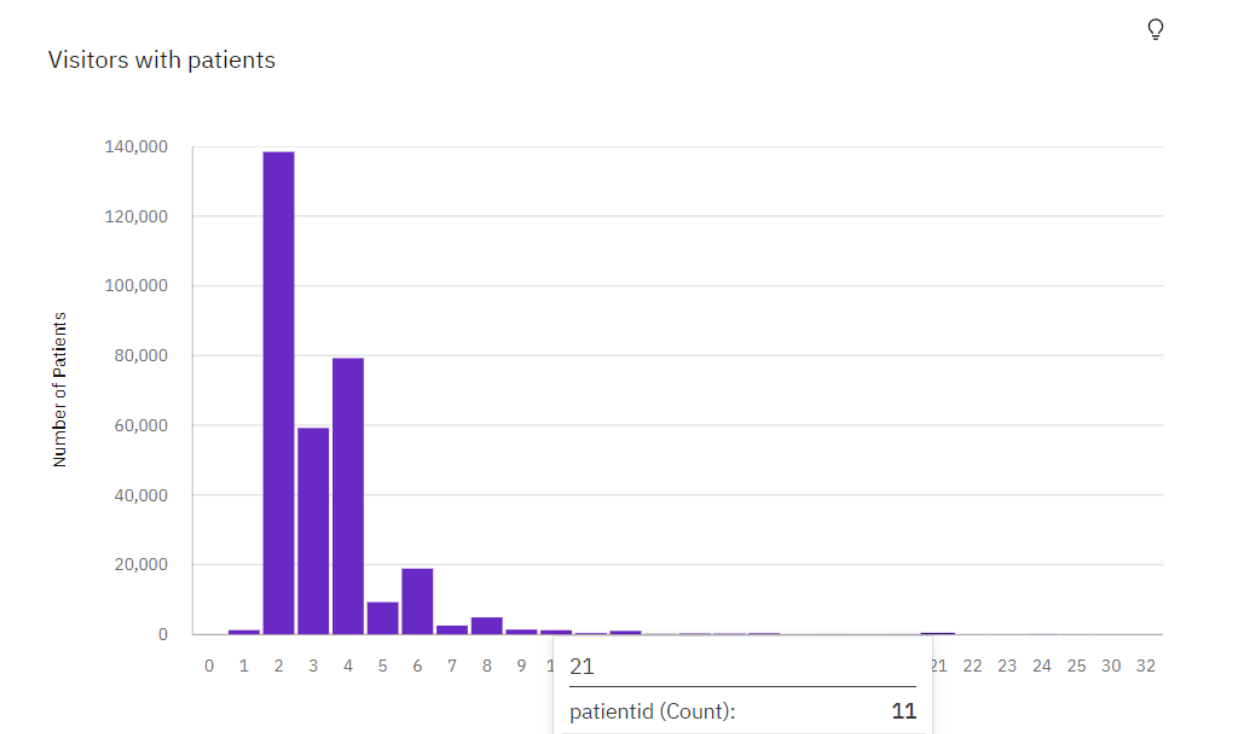
Age distribution



Severity of illness:

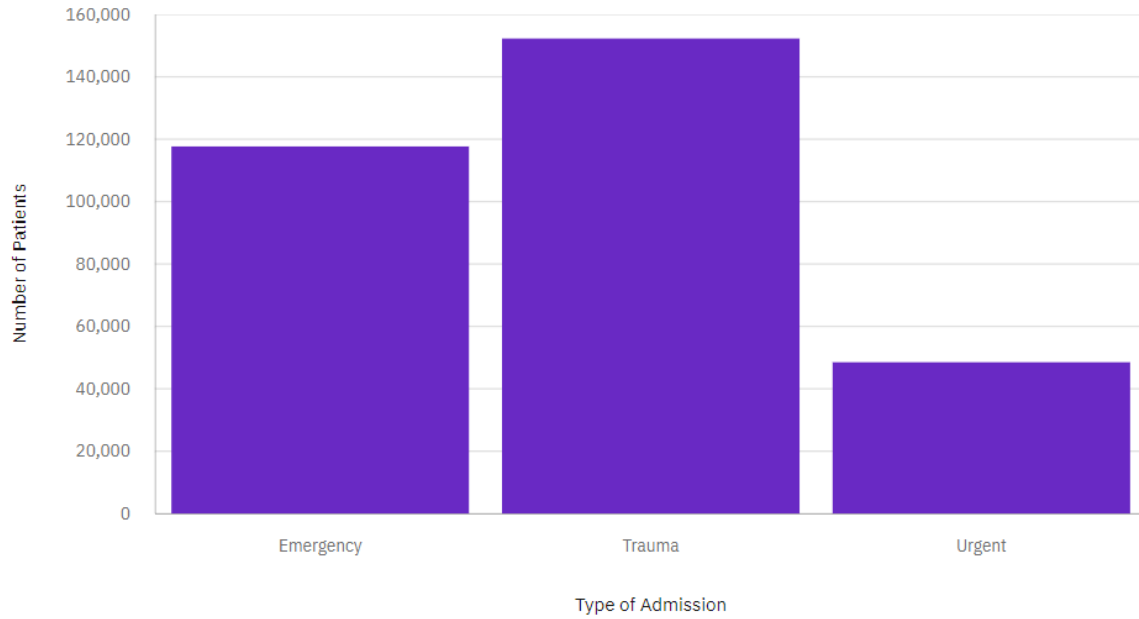


Visitors with patient:



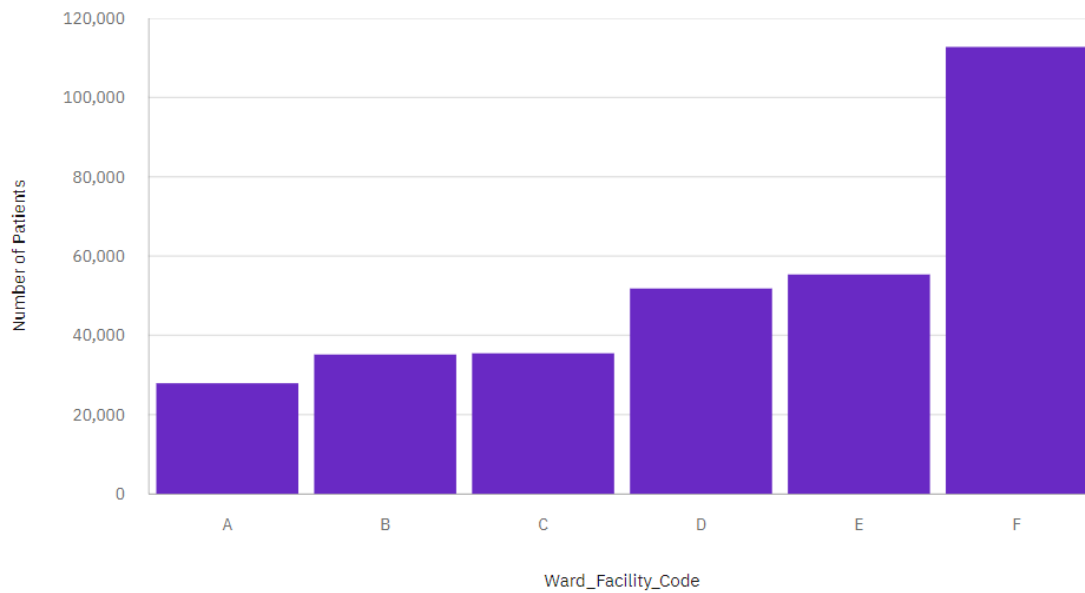
Type of Admission:

Type of admission



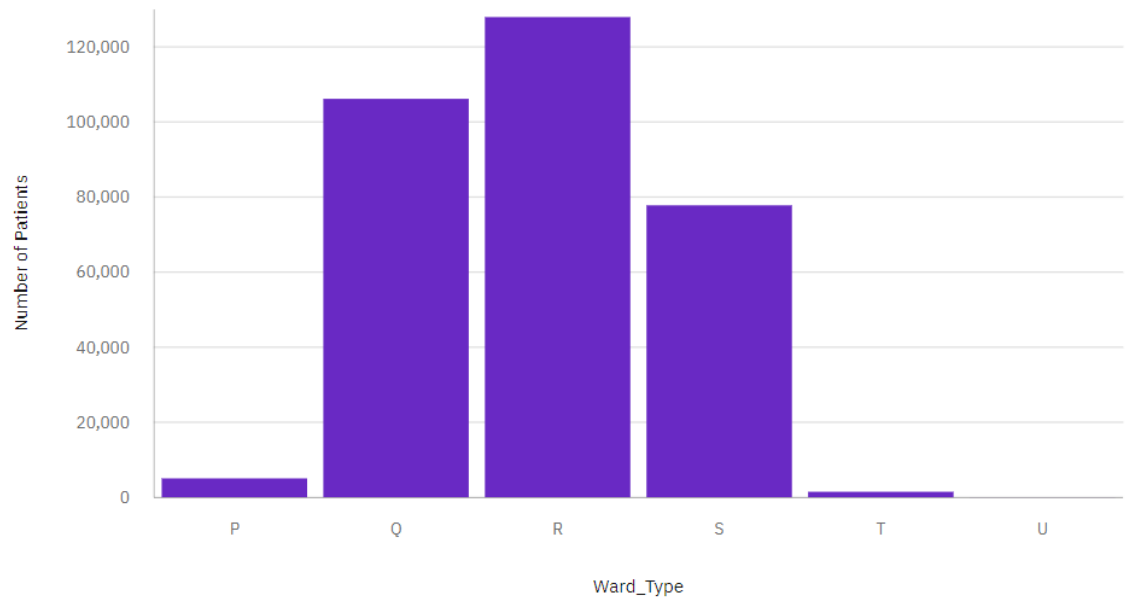
Ward facility code:

Ward facility code



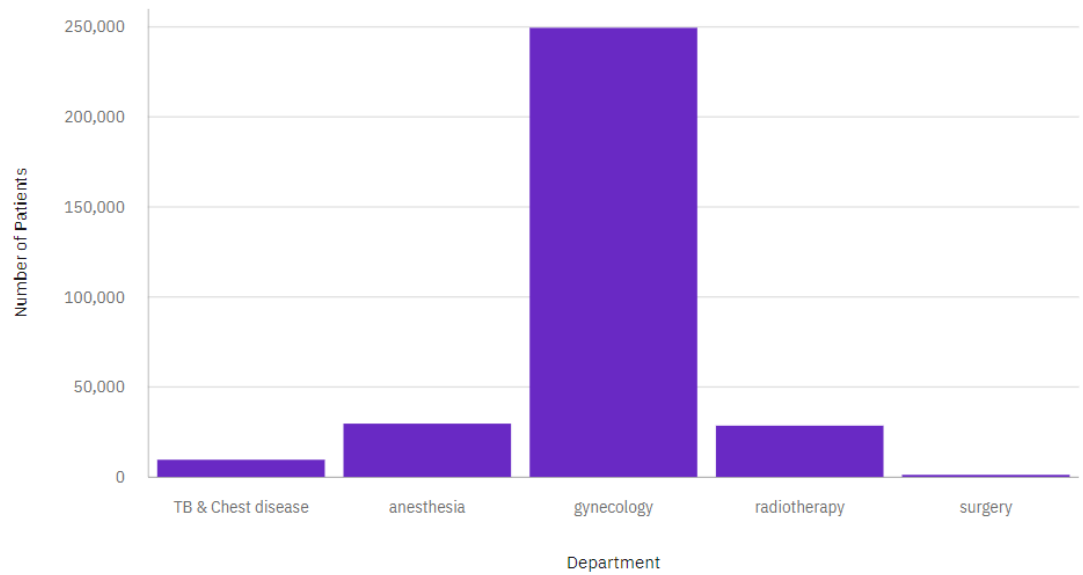
Ward type:

Ward type



Department:

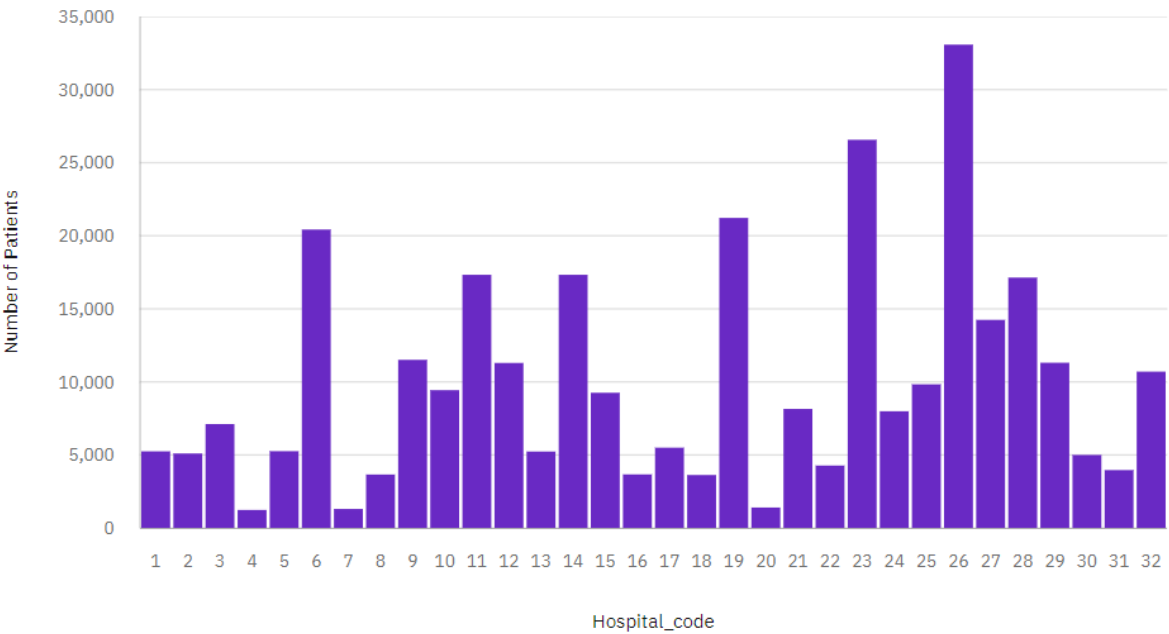
Department



Hospital code:



Hospital code



Hospital type code:



Hospital type code

