Dataset Variables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **Variable name** | **Variable explanation** | **Options** |
| Recipient | 1 | Age |  | 15-77 |
|  | 2 | Gender |  | Male/female |
|  | 3 | BMI | Body-Mass Index | 14-69 |
|  | 4 | Diabetes mellitus |  | Yes/no |
|  | 5 | Arterial hypertension |  | Yes/no |
|  | 6 | Dialysis at transplant |  | Yes/no |
|  | 7 | Etiology |  | 0: virus C cirrhosis  1: alcohol  2: virus B cirrhosis  3: fulminant hepatic failure  4: primary biliary cirrhosis  5: primary sclerosing cholangitis  6: other |
|  | 8 | Portal vein thrombosis |  | 0: no  1: partial  2: complete |
|  | 9 | Days on waiting list |  | 0 - 1078 |
|  | 10 | MELD at inclusion | Model for End-Stage Liver Disease | 1-47 |
|  | 11 | MELD at transplant |  | 5-51 |
|  | 12 | TIPS at transplant |  | yes/no |
|  | 13 | Hepatorenal syndrome |  | yes/no |
|  | 14 | Upper abdominal surgery |  | yes/no |
|  | 15 | Pretransplant status performance |  | 0: at home  1: hospitalized  2: hospitalized in ICU  3: hospitalized in ICU with mechanical ventilation |
|  | 16 | CMV status | Cytomegalovirus | yes/no |
| Donor | 1 | Age |  | 10-87 |
|  | 2 | Gender |  | Male/female |
|  | 3 | BMI | Body-Mass Index | 14-54 |
|  | 4 | Diabetes mellitus |  | Yes/no |
|  | 5 | Arterial hypertension |  | Yes/no |
|  | 6 | Cause of death |  | 0: brain trauma  1: cerebral vascular accident  2: anoxia  3: deceased vascular after cardiac death  4: other |
|  | 7 | Days on ICU |  | 0-59 |
|  | 8 | Hypotension episodes |  | Yes/no |
|  | 9 | High inotropic drug use |  | Yes/no |
|  | 10 | Creatinine plasma level |  | 0.10-9.52 |
|  | 11 | Sodium plasma level |  | 98-188 |
|  | 12 | AST | Aspartate transaminase level | 1-1090 |
|  | 13 | ALT | Alanine aminotransferase plasma level | 2-1400 |
|  | 14 | Total bilirubin |  | 0.06-4.2 |
|  | 15 | Hepatitis B |  | yes/no |
|  | 16 | Hepatitis C |  | yes/no |
|  | 17 | CMV status | Cytomegalovirus | Yes/no |
| Transplant | 1 | Multi-organ | multi-organ harvesting | yes/no |
|  | 2 | Combined transplant |  | yes/no |
|  | 3 | Complete or partial graft |  | yes/no |
|  | 4 | Cold ischemia time |  | 0: <6h  1: 6-12h  2: >12h |
|  | 5 | AB0 compatible transplant |  | yes/no |
| Result | 1 | Graft lost at 3 months |  | yes/no |
| Days | 1 | Survival days after transplant |  |  |

Data preprocessing

1 original row:

|  |  |
| --- | --- |
| etiqueta | AC-001 |
| hospital | AC |
| edadentx | 54 |
| sexorec | 0 |
| bmibasal | 35.08302984 |
| diabetesPreTx | 0 |
| htabasal | 0 |
| dialisis | 0 |
| etiologiaprincipal | 2 |
| trombosisportal | 0 |
| tiempolistaespera | 64 |
| meldinclusion | 12 |
| meldtx | 13 |
| tips | 0 |
| sindromehepatorrenal | 0 |
| apcirugiaabdosuperior | 0 |
| sfiptx | 0 |
| cmvbasal | 1 |
| edaddon | 31 |
| sexodon | 0 |
| bmiestdon | 31.14 |
| diabetesmelitusdon | 0 |
| htadon | 0 |
| causaexitus | 0 |
| diasuci | 18 |
| hipotension | 0 |
| inotropos | 1 |
| creatinina | 1.2 |
| na | 140 |
| ast | 184 |
| alt | 150 |
| bit | 0.9 |
| antihbc | 0 |
| vhc | 0 |
| cmvdon | 1 |
| multiorganico | 1 |
| txcombinado | 0 |
| injertocompletoparcial | 0 |
| tiempoisquemiafria | 1 |
| compatibilidadabo | 1 |
| perdidainjerto3meses | 0 |
| etiqueta | AC-001 |
| dias | 360 |

# Changes in OpenRefine

|  |
| --- |
| Remove column etiqueta |
| Remove column hospital |
| Rename column edadentx to age |
| Rename column sexorec to gender |
| Text transform on cells in column gender using expression value.toNumber() |
| Rename column bmibasal to bmi\_rec |
| Rename column bmi\_rec to bmi |
| Rename column diabetesPreTx to diabetes |
| Rename column htabasal to hypertension |
| Rename column dialisis to dialysis |
| Rename column etiologiaprincipal to etiology |
| Rename column trombosisportal to portal\_thrombosis |
| Rename column tiempolistaespera to waitinglisttime |
| Rename column sindromehepatorrenal to hepatorrenalsyndrome |
| Rename column apcirugiaabdosuperior to upperadbsurgery |
| Rename column sfiptx to pretxstatusperf |
| Rename column cmvbasal to cmv |
| Rename column edaddon to age\_donor |
| Rename column sexodon to gender\_donor |
| Rename column bmiestdon to bmi\_donor |
| Rename column diabetesmelitusdon to diabetes\_donor |
| Rename column htadon to hypertension\_donor |
| Rename column causaexitus to deathcause |
| Rename column diasuci to icudays |
| Rename column inotropos to inotropes |
| Rename column creatinina to creatinine |
| Rename column bit to bilirrubin |
| Rename column antihbc to hep\_b |
| Rename column vhc to hep\_c |
| Rename column cmvdon to cmv\_don |
| Rename column multiorganico to multiorgan |
| Rename column txcombinado to combinedtx |
| Rename column injertocompletoparcial to complete\_parcial\_graft |
| Rename column tiempoisquemiafria to coldischemiatime |
| Rename column compatibilidadabo to ab0comp |
| Remove column etiqueta2 |

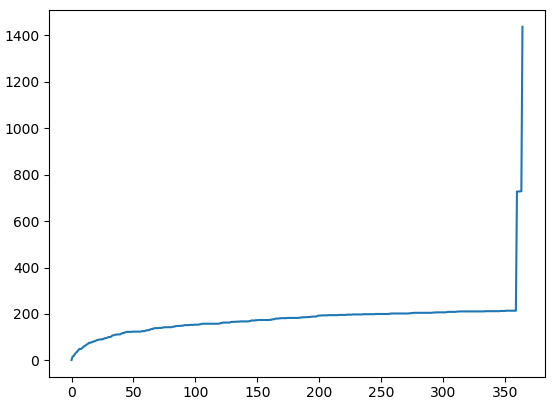
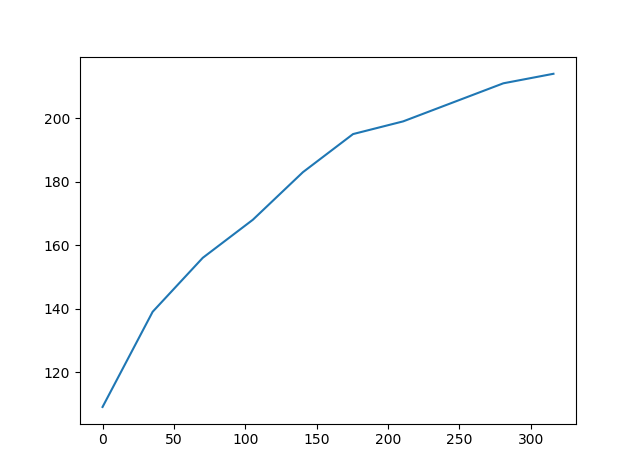
# General info about dataset

# 

# 

# 

# Cumulative frequency table days survived

Changed survived days from a number to categories

|  |  |  |
| --- | --- | --- |
| 1 | before 20 days | 82 |
| 2 | between 20 days and 3 months | 67 |
| 3 | between 3 months and a year | 65 |
| 4 | no graft failure presented after the first year | 1222 |

Imbalanced Ratio: failure – non-failure {214 - 1222}

# Data preprocessing

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Portal thrombosis |  | 0: no | 1: partial | 2: complete |
| 0 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 |
| 2 | 0 | 0 | 1 |

Number of columns equal to the number of categories. Each column corresponds to one type of portal thrombosis and there is either a 1 or a 0. For example, there will be a 1 in the ‘no’ column if there was no portal thrombosis.