$_{file})startTime = datetime.now()$
$_{c}olumns =$
2000
$_{p}ath =$
$config.FIGURES_{D}IRtables_{p}ath =$
$config.TABLES_{D}IRfilename =$
$config.PROCESSED_{F}INAL_{D}ESCRIPTIVEprint(”Loadingdatafile :$
$”, filename)data =$
$pd.read_{p}ickle(filename)print(”Loaded.”)print(list(data))$
$_{p}ts =$
$data[”patientid”].nunique()n_{e}ncs =$
$data[”encounterid”].nunique()n_{3}0dreadmits =$
$data[”readmitted30d”].sum()readmit_{r}ate =$
$n_{3}0dreadmits/n_{e}ncs$
$_{m}edians =$
$data.groupby([’readmitted30d’])[”patient_{a}ge”].median()age_{m}edian_{n}onreadmit =$
$age_{m}edians.iloc[0]age_{m}edian_{r}eadmit =$
$age_{m}edians.iloc[1]$
$_{r}eadmission =$
$len(blackdf[blackdf[”readmitted30d”] ==$
$1])/len(blackdf)$
$_{r}eadmission =$
$len(whitedf[whitedf[”readmitted30d”] ==$
$1])/len(whitedf)$
$_{s}tatus”] ==$
$”Marriedorpartnered”]marrieddf =$
$marrieddf[[”marital_{s}tatus”, ”readmitted30d”]]married_{r}eadmission =$
$len(marrieddf[marrieddf[”readmitted30d”] ==$
$1])/len(marrieddf)$
$_{s}tatus”] ==$
$”Divorcedorseparated”]divorceddf =$
$divorceddf[[”marital_{s}tatus”, ”readmitted30d”]]divorced_{r}eadmission =$
$len(divorceddf[divorceddf[”readmitted30d”] ==$
$1])/len(divorceddf)$
$_{s}tatus”] ==$
$”Widowed”]widoweddf =$
$widoweddf[[”marital_{s}tatus”, ”readmitted30d”]]widowed_{r}eadmission =$
$len(widoweddf[widoweddf[”readmitted30d”] ==$
$1])/len(widoweddf)$
$_{s}tatus”] ==$
$”Single”]singledf =$
$singledf[[”marital_{s}tatus”, ”readmitted30d”]]single_{r}eadmission =$
$len(singledf[singledf[”readmitted30d”] ==$
$1])/len(singledf)$
$_{M}edicare”] ==$
$1]medicare =$
$medicare[[”financialclass_{M}edicare”, ”readmitted30d”]]medicare_{r}eadmission =$
$len(medicare[medicare[”readmitted30d”] ==$
$1])/len(medicare)$
$_{P}rivate_{H}ealth_{I}nsurance”] ==$
$1]private =$
$private[[”financialclass_{P}rivate_{H}ealth_{I}nsurance”, ”readmitted30d”]]private_{r}eadmission =$
$len(private[private[”readmitted30d”] ==$
$1])/len(private)$
$_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.min()mean_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.mean()q1_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.quantile(0.25)median_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.median()q3_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.quantile(0.75)max_{l}os_{p}er_{p}t =$
$data.length_{o}f_{s}tay_{i}n_{d}ays.max()$
$_{o}f_{s}tay_{i}n_{d}ays =$
$data0.length_{o}f_{s}tay_{i}n_{d}ays.apply(np.floor)rounddownmode_{l}os_{p}er_{p}t =$
$data0.length_{o}f_{s}tay_{i}n_{d}ays.mode()$
$_{e}nc =$
$data[[”patientid”, ”encounterid”]]df =$
$(pt_{e}nc.groupby(”patientid”)[”encounterid”].nunique().sort_{v}alues(ascending =$
$False).reset_{i}ndex(name =$
$”encounternum”))$
$_{a}dmit =$
$min(data[”admissiontime”])latest_{a}dmit =$
$max(data[”admissiontime”])earliest_{d}ischarge =$
$min(data[”dischargetime”])latest_{d}ischarge =$
$max(data[”dischargetime”])$
$_{e}nc_{p}er_{p}t =$
$df.encounternum.min()mean_{e}nc_{p}er_{p}t =$

$_{d}f.head())print("Makingcsvwithimportantnumbers...")csv_{d}f.to_{c}sv(config.PAPER_{N}UMBERS, m$
"w", header =
True)
$_{e}ncs$:
, hospitalizationsforn$_{p}$ts :,uniquepatients, pts$_{o}$ne$_{e}$ncounter :,(pts$_{o}$ne$_{e}$ncounter/n$_{p}$ts $*$ 100 : .0fsentence0
f"Themediannumberofhospitalizationsperpatientwasmedian$_{e}$nc$_{p}$er$_{p}$t : .0f(rangemin$_{e}$nc$_{p}$er$_{p}$t : .0f−
max$_{e}$nc$_{p}$er$_{p}$t : .0f, [q1$_{e}$nc$_{p}$er$_{p}$t, q3$_{e}$nc$_{p}$er$_{p}$t])."sentence03 =
f"Thereweren$_{3}$0readmits :,thirty−
dayunplannedreadmissionsforanoverallreadmissionrateofreadmit$_{r}$ate $*$ 100 : .0fsentence04 =
f"Themedianlengthofstaywasmedian$_{l}$os$_{p}$er$_{p}$t :, .2fdays(rangemin$_{l}$os$_{p}$er$_{p}$t : .0f−
max$_{l}$os$_{p}$er$_{p}$t : .0f, [q1$_{l}$os$_{p}$er$_{p}$t, q3$_{l}$os$_{p}$er$_{p}$t])."sentence05 =
f"Thedemographicandclinicalcharacteristicsofthepatientcohortaresummarizedin Table1."sentence06a =
f"Higherratesofunplanned30−
dayreadmissionswereobservedinpatientswhowereolder(medianageage$_{m}$edian$_{r}$eadmit : .0fvs.age$_{m}$edian$_{r}$
f"black(rateofblack$_{r}$eadmission $*$ 100 : .0fsentence06c =
f"divorced/separatedorwidowed(ratesofdivorced$_{r}$eadmission $*$ 100 : .0fsentence06d =
f"onMedicareinsurance(rateofmedicare$_{r}$eadmission $*$ 100 : .0fsentence06e =
f"andhadconditionssuchascancer, renaldisease, congestiveheartfailure, andCOPD(Table1)."paragraph
sentence01+
sentence02+
sentence03+
sentence04+
sentence05+
sentence06a+
sentence06b+
sentence06c+
sentence06d+
sentence06e
$_{D}IR$) :
print("Makingfoldercalled", config.TEXT$_{D}$IR)os.makedirs(config.TEXT$_{D}$IR)
$_{t}ext_{f}ile$ =
config.TEXT$_{D}$IR/"results$_{p}$aragraphs.txt"...andsave.withopen(results$_{t}ext_{f}ile$, "w")astext$_{f}ile$:
print(paragraph01, file =
text$_{f}ile$)
$_{f}ile_{l}atex$ =
config.TEXT$_{D}$IR/"results$_{p}$aragraphs$_{l}$atex.txt"andmakeaLaTeX−
friendlyversion(escapetheReadinthefilewithopen(results$_{t}ext_{f}ile$,' r')asfile :
filedata =
file.read()Replacethetargetstringfiledata =
filedata.replace('Writethefilewithopen(text$_{f}ile_{l}atex$,' w')asfile :
file.write(filedata)
$_{c}lf_{d}f$ =
pd.read$_{c}$sv(config.TRAINING$_{R}$EPORTS)results$_{r}eg_{d}f$ =
pd.read$_{c}$sv(config.REGRESSOR$_{T}$RAINING$_{R}$EPORTS)
$_{c}lf_{d}f$ =
results$_{c}lf_{d}f$[results$_{c}lf_{d}f$.Time! =
"Time"]maketheTimecolumnapandasdatetimecolumnresults$_{c}lf_{d}f$['Time'] =
pd.to$_{d}$atetime(results$_{c}lf_{d}f$['Time'], format =
"results$_{c}lf_{d}f$ =
results$_{c}lf_{d}f$[["Time","Target","ROCAUC"]]print(results$_{c}lf_{d}f$)print(results$_{c}lf_{d}f$[["Time","Target","
pd.to$_{n}$umeric(results$_{c}lf_{d}f$["ROCAUC"])selecttherunwiththefullnumberoffeatures?results$_{c}lf_{d}f$ =
results$_{c}lf_{d}f$.loc[results$_{c}lf_{d}f$.groupby('Target')["NumberofFeatures"]idxmax()]selectthenewesttraining
results$_{c}lf_{d}f$.loc[results$_{c}lf_{d}f$.groupby('Target').Time.idxmax()]selecttherunwithhighestAUCresults$_{c}lf_{d}$
results$_{c}lf_{d}f$.loc[results$_{c}lf_{d}f$.groupby('Target')["ROCAUC"].idxmax()]print(results$_{c}lf_{d}f$)
$_{T}EX$,' w')astf :
tf.write(results$_{c}lf_{d}f$.to$_{l}$atex())