Thesis SAS CODES

```
libname TMP1 "H:\ijoof"; RUN;
proc contents data= tmp1.seerprost;run; * looking at Covariates;
*Covert these var from char to numeric;
data TMP;
set tmp1.seerprost;
age = input(Age_at_diagnosis,best12.);
income= input(MedianfamilyincomeintensACS2010,best12.);
survival = input(Survival_months,best12.);
diagnosis = input(Year_of_diagnosis,best12.);
run;
*Checking whether we transfer these variable from char to num;
proc contents data=TMP;run;
proc print data=TMP (obs=10);run;
*Select a subdataset for those whose ags is from 45 to 69;
data TMP1;
set TMP;
where age >=45 and age <= 69;
run;
*Checking the results;
proc univariate data=tmp1;
var age;
run;
*******Age**********;
*incarceration average using median;
proc univariate data=TMP1;
var Incarceration_rate;
run;   *median =0.0422;
*Age with high incarceration, STD = 5.78 and median = 62;
proc univariate data = TMP1;
var age;
where Incarceration_rate > 0.042200;
run;
*Age with low incarceration, std is = 5.76, median = 62;
proc univariate data = TMP1;
var age;
where Incarceration_rate < 0.042200;
run;

*******survival*********;
*surviva with high incarceration;
data tmp1;
set tmp1;
if survival=. then delete;run;
proc univariate data = TMP1;
var survival;
where Incarceration_rate > 0.042200;
run;
*surviva with low incarceration;
proc univariate data = TMP1;
var survival;
where Incarceration_rate < 0.042200;
run;

proc freq data=TMP1;
```

```sas
       tables SEERcausespecificdeathclassific;
       run;
       ********survival with outcome (death)*********;
       proc univariate data = TMP1;
       var survival;
       where Incarceration_rate > 0.042200 and
       SEERcausespecificdeathclassific="Dead (attributable to this can";
       run;
       proc univariate data = TMP1 ;
       var survival;
       where Incarceration_rate < 0.042200 and
       SEERcausespecificdeathclassific="Dead (attributable to this can";
       run;

       *other covariates;
       proc freq data = tmp1;
       tables RaceandoriginrecodeNHWNHBNHAIAN;run;
       proc freq data= TMP1;
       tables RaceandoriginrecodeNHWNHBNHAIAN;* race;
       where Incarceration_rate > 0.042200;run;
       proc freq data= TMP1;
       tables RaceandoriginrecodeNHWNHBNHAIAN;
       where Incarceration_rate < 0.042200;run;

       * Grade;
       proc freq data= TMP1;
       tables Grade;
       where Incarceration_rate > 0.042200;run;
       proc freq data= TMP1;
       tables Grade;
       where Incarceration_rate < 0.042200;run;

       * stage;
       proc freq data= TMP1;
       tables Summary_stage_2000_1998;
       where Incarceration_rate > 0.042200;run;
       proc freq data= TMP1;
       tables Summary_stage_2000_1998;
       where Incarceration_rate < 0.042200;run;

       *Chemo ;
       proc freq data= TMP1;
       tables Chemotherapy_recode_yes_no_unk;
       where Incarceration_rate > 0.042200;run;
       proc freq data= TMP1;
       tables Chemotherapy_recode_yes_no_unk;
       where Incarceration_rate < 0.042200;run;


       **********radiation**********;
       proc freq data= TMP1; *recode radiation in 3 groups;
       tables Radiation_recode/MISSING;
       run;

       *recode radiation in 3 groups;
       *I fixed the missing value problem here, no worries anymore;
       data TMP1;
```

```sas
set TMP1;
if Radiation_recode= "Beam radiation"
or Radiation_recode= "Combination of beam with implants or isotopes"
or Radiation_recode= "Radiation, NOS method or source not specified"
or Radiation_recode= "Radioisotopes (1988+)"
or Radiation_recode= "Radioactive implants (includes brachytherapy) (1988+)"
then radiation= 1;

else if Radiation_recode= "Recommended, unknown if administered"
or Radiation_recode= "Refused (1988+)"
then radiation = 2;

else radiation = 3;
run;

proc freq data= TMP1;
tables radiation;
where Incarceration_rate > 0.042200;run;
proc freq data= TMP1;
tables radiation;
where Incarceration_rate < 0.042200;run;


*create income cat;
data TMP1;
set TMP1;
income = income * 10;
run;
proc univariate data=TMP1;
var income;
run;
data TMP1;
set TMP1;
if income = . then med_income = .;
else if income < 30000 then med_income = 1;
else if income < 34999 & income >=30000 then med_income =2;
else if income < 46000 & income >=34999 then med_income =3;
else med_income=4;
run;
proc univariate data = TMP1; * mean income  is 37777;
var med_income;run;

proc freq data= TMP1;
tables med_income;
where Incarceration_rate > 0.042200;run;
proc freq data= TMP1;
tables med_income;
where Incarceration_rate < 0.042200;run;

*we cannot make that;
proc contents data=tmp1;run;
proc print data=tmp1 (obs=10);var VHighschooleducationACS_2010_14; run;
proc freq data=tmp1;
tables VHighschooleducationACS_2010_14;
tables V9thgrade_education_ACS_2010_14;
run;
*recode edu var into two categories;
```

```sas
*Making status var, status is the outcomedeath in your data;

data TMP1;

set TMP1;

if SEERcausespecificdeathclassific = "Dead (attributable to this can" then
status = 1;

else status = 0;

run;

proc freq data=TMP1; *check;

tables status;

run;

* Process the incarceration;

data TMP1;

set TMP1;

if Incarceration_rate>=0.0422 then incar = 1;

else incar = 0;

run;

proc freq data=TMP1;

tables incar;

run;

*Process race;

data TMP1;

set TMP1;

if RaceandoriginrecodeNHWNHBNHAIAN = "Non-Hispanic Black" then race = 1;

else race = 0;

run;


*Survival analysis model;


*contrasting by race when incar is high or low table 2a;
```

```sas
proc phreg data = tmp1;
model survival*Status(0)= Incar age race  incar*race/rl;
contrast "Black Vs White low"  race 1 incar*race 0 /estimate = exp;* contrast
amoung Low stratified hazard ratio obtained from point estimate;
contrast "Black Vs White high"  race 1 incar*race 1 /estimate = exp;
hazardratio incar; *hazard ratio for Incar overall and CI;
run;


*median income;

proc phreg data=TMP1;

model survival*status(0) = incar age race med_income med_income*race /RL;

contrast "Black Vs White" med_income 1 med_income*race 1 / estimate=exp;

contrast "black Vs White" med_income 2 med_income*race 2 / estimate=exp;
contrast "black Vs White" med_income 3 med_income*race 3 / estimate=exp;
contrast "black Vs White" med_income 4 med_income*race 4 / estimate=exp;
hazardratio med_income ; *overal hazard ratio;
run;


*table 2b multivariable adjusted;

proc phreg data = tmp1;
model survival*Status(0)= Incar age med_income race grade_cat stage
incar*race/rl;
contrast "Black Vs White low"  race 1 incar*race 0 /estimate = exp;* low
incarcerationstratified hazard ratio obtained from point estimate;
contrast "Black Vs White high"  race 1 incar*race 1 /estimate = exp;*high
incaceration;
hazardratio incar; *hazard ratio for Incar overall and CI;
run;
proc freq data=TMP1;

where incar<0.0422;

tables status;  * finding the N deaths;

run;

proc freq data=TMP1;

where incar >0.0422;

tables status;

run;



proc freq data=TMP1;
```

```
    where incar<0.0422 and race=1;   *finding death Numbers for NHB,subtracted to
    get numbers for NHW;

    tables status;

    run;

    proc freq data=TMP1;

    where incar>0.0422 and race=1;

    tables status;

    run;


    proc freq data=TMP1;

    where med_income=1;

    tables status;  * finding the N deaths for median income;

    run;

    proc freq data=TMP1;

    where med_income=2;

    tables status;

    run;

    proc freq data=TMP1;

    where med_income=3;

    tables status;

    run;

    proc freq data=TMP1;

    where med_income=4;  * finding the N deaths for median income;

    tables status;

    run;


    proc freq data=TMP1;

    where med_income=1 and race=1;

    tables status;  * finding the N deaths for median income by race;

    run;
```

```sas
proc freq data=TMP1;

where med_income=2 and race=1;

tables status;

run;

proc freq data=TMP1;

where med_income=3 and race=1;

tables status;

run;

proc freq data=TMP1;

where med_income=4 and race=1;  * finding the N deaths for median income by race;

tables status;

run;
*table 3;

proc print data= TMP1 (obs=2);run;
data TMP1;
SET TMP1;
if Summary_stage_2000_1998 = "Localized" then stage = 1;
else if Summary_stage_2000_1998 = "Regional" then stage =2;
else if Summary_stage_2000_1998 = "Distant" then stage =3;
else stage=.;
run;




****** Grade******;
proc freq data = TMP1;
tables grade;
run;
data TMP1;
set TMP1;
if grade="Well differentiated; Grade I" then grade_cat = 1;
else if grade="Moderately differentiated; Grade II" then grade_cat = 2;
else if grade = "Poorly differentiated; Grade III" then grade_cat = 3;
else grade_cat = .;
run;
proc freq data = TMP1;
tables grade_cat;
run;


*
*Multivariable adjusted OR 3a;
```

```
*using  logistic regression;



*score p-value <0.0001 PO assumption not met?;
*Multivariable stage;
proc logistic data = TMP1;
model stage(Ref = '1') = incar age Med_income / link = glogit;
 *table 3a;
run;


*PLR;
proc logistic data = TMP1;

model grade_cat(ref= '1') = incar age Med_income / link = glogit;
 *table 3a;
run;


*Effect of survial staus on race;

proc phreg data = tmp1;
model survival*Status(0)= race age/rl;
contrast "Black Vs White low"  race 1  /estimate = exp; run;
```

Table 3. Multivariable-adjusted odds ratios estimating the association between incarcerati

	Stratified Effects OR (95% CI) high vs. low incarceration
Stage	
I	Referent
II	0.79(0.74-0.84)
III	1.18(1.07-1.31)
Tumor differentiation	
Well differentiated	Referent
Moderately	1.06(0.99-1.14)
Undifferentiated	1.10(1.03-1.19)

Table 3: Age adjusted odds ratios (OR) and 95% confidence intervals for the association between incarceration rates and tumor characteristics

Incarceration	Stratified Effects OR		
	Stage I	Stage II	Stage III
Low	referent	referent	referent
High	1.08(0.78- 1.49)	1.18(0.68- 2.04)	1.29(0.59- 2.81)
	Well differentiated	Moderately differentiated	Undifferentiated
Low	referent	referent	referent
High	0.90(0.63-1.28)	0.82(0.39- 1.71)	0.74(0.22- 2.45)