

```
*****;
*Programmed By: Thomas Grover
*Programmed To: Analyze NC Senior Games Data
*****;
```

```
*****;
*Section One: Setting up our workplace and Importing Necessary Data Files, Creating format ;
*****;
```

```
libname PBProj '/home/u63735775/PickleballProject';
filename reffile '/home/u63735775/PickleballProject/NCSGPickleballData.xlsx';
```

```
*NOTE: Commented out because we only need to import once;
/* proc import datafile = reffile */
/*      dbms = XLSX */
/*      out = work.data; */
/* getnames = yes; */
/* RUN; */
```

```
proc format;
  value $ActivityCategory
    'Pickleball' = 'Pickleball'
    'Badminton', 'Table Tennis', 'Tennis' = 'Racket'
    'Visual Arts', 'Literary Arts', 'Heritage Arts', 'Performing Arts' = 'Arts'
    'Cycling', 'Disc Golf', 'Fun Walk', 'Distance Run', 'Track' = 'Outdoor'
    other = 'Other';
run;
```

```
*****;
*Section Two: Data Preparation (cleaning Raw data, creating the datasets we will need in our analysis ;
*****;
```

```
*Data step to clean up the data, assigning labels to make
things look nicer, dropping unnecessary variables, creating
a new variable named "Era" for pre/post covid records.;
```

```
data PBProj.Data;
  attrib ParticipantID label = "ParticipantID"
    EventType label = "Event Type"
    ActivityCategory label = 'Activity Category'
    EventYear label = "Event Year"
    Era label = "Pre|Post-Covid"
    BirthYear label = "Birth Year"
    Age label = "Participant Age"
    AgeGroup label = "Participant Age Group";
  set Work.Data(rename=(EventYearParticipantAge = Age
    EventYearParticipantAge_1 = AgeGroup)
    drop = H I);
  if EventYear ge 2020 then Era = "Post-Covid";
  else Era = "Pre-Covid";
  ActivityCategory = put(EventType, $ActivityCategory.);
run;
```

```
*Creating our pickleball dataset;
proc sort data = PBProj.Data
  out = PBProj.Pickleball;
  by descending EventYear;
  where EventType = "Pickleball";
run;
```

```
*Creating our pickleball participation count/year dataset;
```

```

proc freq data = PBProj.Pickleball;
  by Era;
  tables EventYear / nocum out=PBProj.PickleballCounts;
run;

*Creating our age dataset;
proc sort data = PBproj.data
  out = PBproj.Age(drop = EventType EventYear BirthYear ParticipantID)
  nodupkey;
  by ParticipantID;
run;

proc sort data = PBProj.Age;
  by Era;
run;

*****;
*Section Three: Data Analysis (analyzing participation frequency ;
*across the years, age and demographic information for participants;
*****;

title "Frequencies of Pickleball Participation by Year";
footnote j=left "If a player participated multiple times, they are counted multiple times";
proc report data = PBProj.PickleballCounts;
  columns Era EventYear Count;
  define Era / group
    order = formatted descending
    style = {fontstyle=italic}
    'Pre|Post Covid';
  define EventYear / group order = freq;
  define Count / analysis 'Count'
    style = {fontweight=bold};
run;
title;
footnote;

proc ttest data = PBProj.PickleballCounts;
  class Era;
  var count;
run;

*Plotting the count for each year over time, seperating
the eras, in order to visualize the trend of pickleball;
title "Participant Count Over Time by Era";
proc sgplot data=PBProj.PickleballCounts;
  vbar EventYear / response=count
    group=era
    groupdisplay=cluster
    datalabel
    datalabelattrs=(size=12pt weight=bold)
    clusterwidth = .95
    dataskin = pressed;
  keylegend / position = nw
    location = inside
    title = ""
    titleattrs = (size=14pt weight=bold)
    valueattrs = (size=14pt);
  xaxis label="Year"
    labelattrs=(size=16pt)
    valueattrs=(size=12pt weight=bold);
  yaxis label="Participant Count"

```

```
labelattrs=(size=16pt)
valueattrs=(size=12pt weight=bold);
run;
title;
```

```
title "Age Analysis with Histogram";
proc univariate data=PBProj.Age
    novarcontents;
    var age;
    histogram age / midpoints = 45 to 100 by 1 normal kernel barwidth=1;
    inset mean median std skewness range n / position=ne;
run;
title;
```