Assignment 1

Please use the Pandas and Numpy libraries to answer the question I posed in the document. You will create a markdown cell with the question in it and then answer the question in the next cell.

1) Load in the appropriate csv file as a pandas dataframe (batting.csv)
2) Print out the dimensions and info about the dataframe you just created

For next part I am going to ask you a series of question, for some I am going to want to simply see a number. For others I will want to see the entire dataframe. I will make it very clear which is which.

3) How many players have hit 40 or more HRs in one single season? (Number only)
4) How many players have hit 600 or more HRs for their career? (Dataframe with name and player first and last name only and amount of HRs)
5) How many unique players have hit 40 2Bs or more, 10 3Bs or more, 200 Hits or more, and 30 HRs or more in one season? (Number Only)
6) How many players seasons have had 100 or more SBs? (Dataframe, first name, last name, yearID, number of SBs, ordered from descending SBs)
7) How many players in the 1960s (yearID) have hit 200 or more HRs? (Dataframe, first name, last name, number of HRs, ordered descending by HR amt)
8) Who has hit the most HRs in history? (Dataframe, first name, last name, HRs
9) Who had the most hits (H) in the 1970s? (Dataframe)
10) Top 5 highest OBP (on base percentage) with at least 500 PAs in 1977? (Dataframe, first name, last name, OBP, ordered descending by OBP)
11) Top 8 highest averages in 2013 with at least 300 PAs? (Dataframe, first name, last name, average, descending by average) -) – batting average, on base percentage calculation link is there.
12) Leaders in hits from 1940 up to and including 1949. (Dataframe, first name, last name, number of hits) (Top 5 is required)
13) Who led MLB with the most hits the most times? And how many times? (Dataframe, Number of hits) – per season
14) Which players have played the most games for their careers? Top 5 first name, last name, descending by games played presented as a dataframe
15) How many players have had more 3000 or more hits for their careers while also hitting 500 or more HRs? Just a number is okay here
16) How many HRs were hit during the entire 1988 season? Just a number is okay here
17) Please filter out and show me the top 3 average seasons by Wade Boggs during his career in seasons in which he had at least 500 or more ABs. Dataframe, first name, last name, average, descending by average
18) Please filter out the top 10 OBPs for the 1995 season with 400 or more PAs, sorted by OBP descending. Dataframe with first name, last name, OBP

19) Who had the most 3Bs (in total) in 1922, 1925, 1926, and 1928? I would like a dataframe with first name, last name, number of 3Bs (only one person is required)
20) How many unique players have hit 30 or more HRs in season while also stealing (SB) 30 more or bases? A number is okay here
21) Who had the highest OBP is 1986 with 400 or more ABs? (Dataframe first name, last name, OBP)
22) Same question but for 1997 and only in the NL (check league ID)? (Dataframe, first name, last name OBP)
23) Who had more than the league average HRs (for players with 500 or more ABs) in 2012 (filter out all players with less 500 ABs)? (Dataframe first name, last name, HR descending) (500 At bats in that league, MLB, ML… does not matter, only year matters)
24) Who is the youngest player to hit 50 or more HRs in a single season? (Dataframe, first name, last name, HRs, season)
25) Who are the five youngest players to hit 300 or more HRs for their career? (Dataframe, first name, last name, season they eclipsed more than 300 HRs)

Grading: Each task will be worth 3.5 pts (there will be 25 total). Naming convention and github link will be worth 6.25 pts apiece. I will include the two bonus questions which will be worth 1 pt apiece (they will be simple viz questions)

BONUS: Graph total HRs per season using bar graph

Using a line graph please graph the average HRs per AB (think about this) per season

Helpful links: https://www.baseball-reference.com/bullpen/On_base_percentage

https://www.baseball-reference.com/bullpen/Batting_average