Put each on its own slide(s) and include your code!

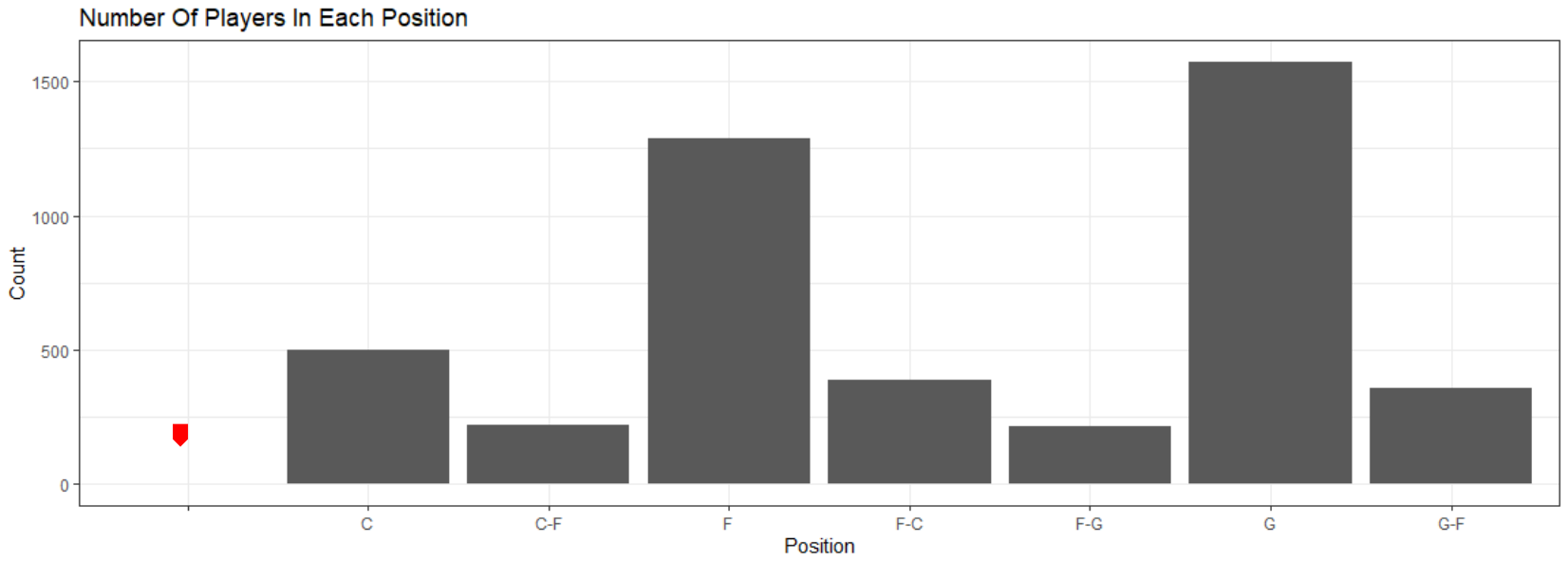
Download into R Studio the PlayerBBall.csv dataset. This data set is every NBA basketball player from 1950 to present. It contains their height, weight, position and the year they played (among other data.) (Position: F-Forward, C-Centers, F-C and C-F – Forward /Centers, G – Guards, F-G – Forward/Guards) FYI: If you feel that these quGGestions are open ended or at least a little vague, this is on purpose. Answer the question as you understand it and make any assumptions you need to make to answer the question and record those assumptions. (3-5 hours)

- Use the PlayerBBall.csv dataset to visually represent (summarize) the number of players in each position.
- Use the dataset to visually investigate the distribution of the weight of centers (C) is greater than the distribution of the weight of forwards (F).
- Use the dataset to visually investigate if the distribution of the height of centers (C) is greater than the distribution of the height of forwards (F).
- Use the dataset to visually investigate if the distribution of height is different between *any* of the positions.
- Use the dataset to investigate how the player's height is related to the player's weight. How does height change as the weight changes?
- Is their any difference in the relationship between height and weight between positions? Are height and weight related differently for different positions.
- A historian would like to investigate the claim that the heights of players have increased over the years. Analyze this claim graphically / visually.
- Create a 3D plot of height vs. weight vs. year and color code the points by position.
- Go to this website and use one of the 50 best plots to visualize some aspect of the data and provide at least one insight. You will present your work in breakout! http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html

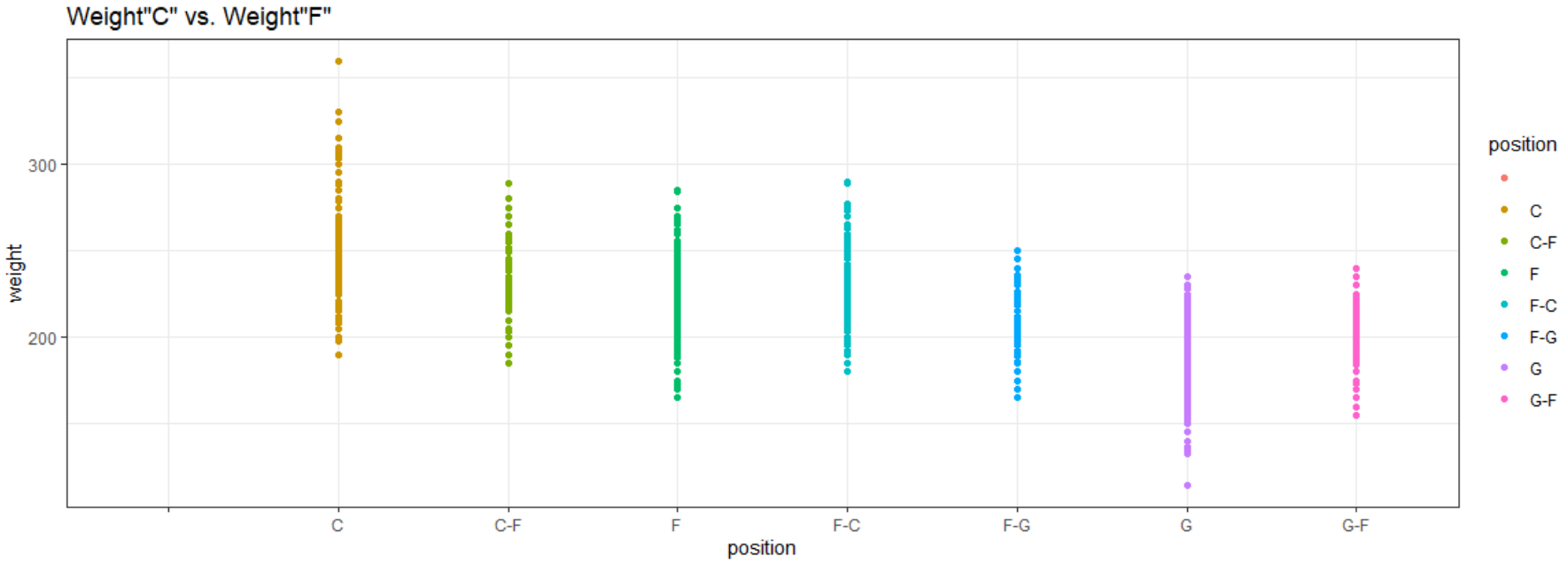
Separate dataset: The EducationIncome.csv dataset has incomes of randomly selected Americans and their level of education. (1-2 hours)

- Visually test the claim that the distribution of incomes increase (mean or median) as the education level rises.

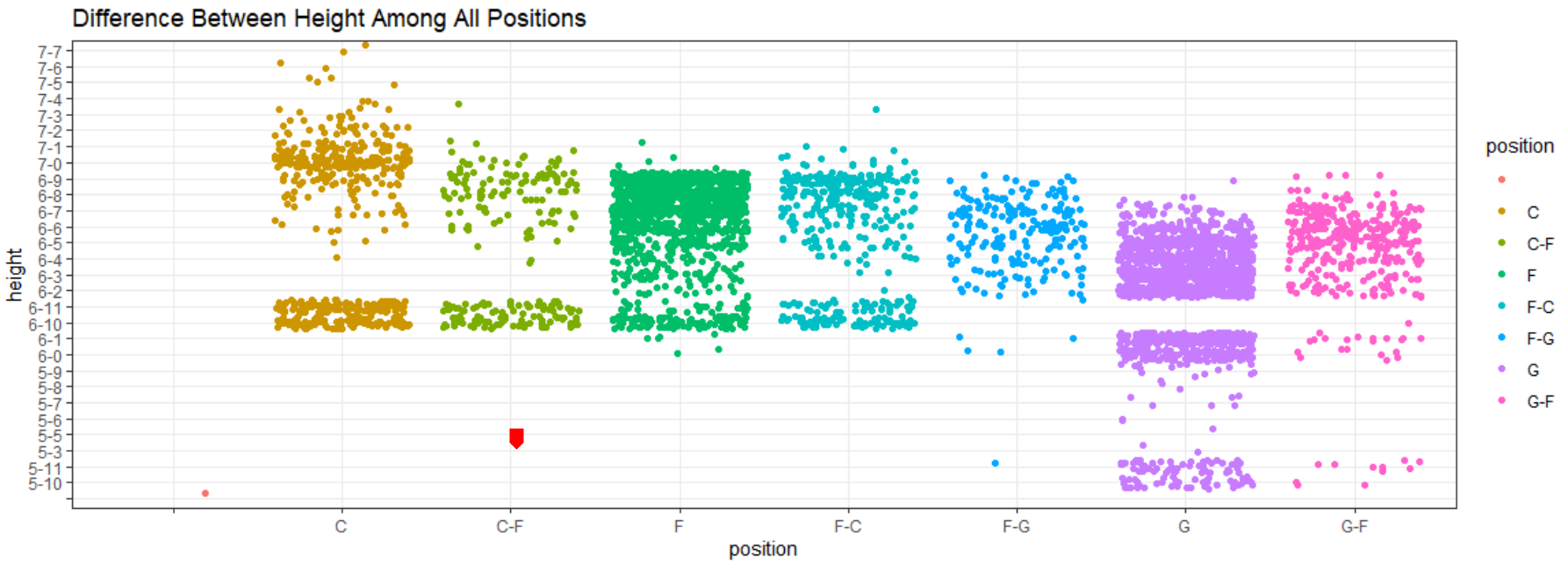
1. Number Of Players In Each Position



2. Weight "C" vs Weight "F"

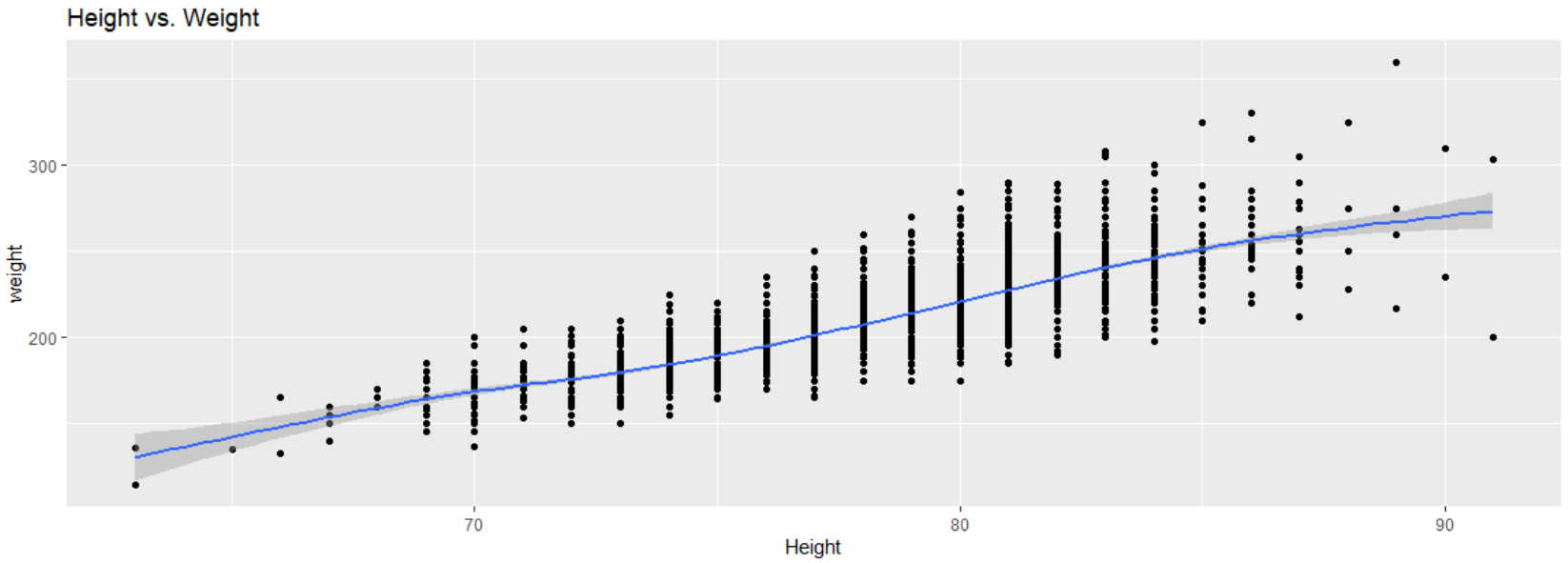


4.Difference Between Height Among All Positions



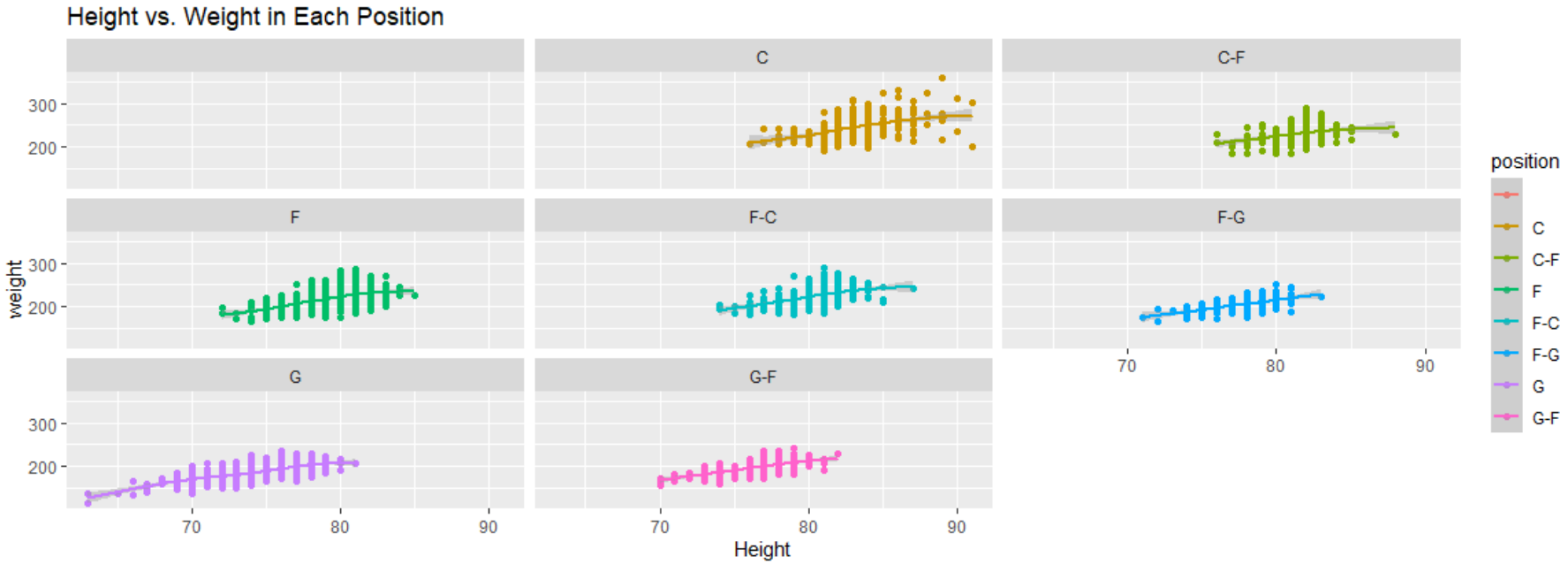
5. Relationship Between Height vs. Weight

As height increases, weight increases.



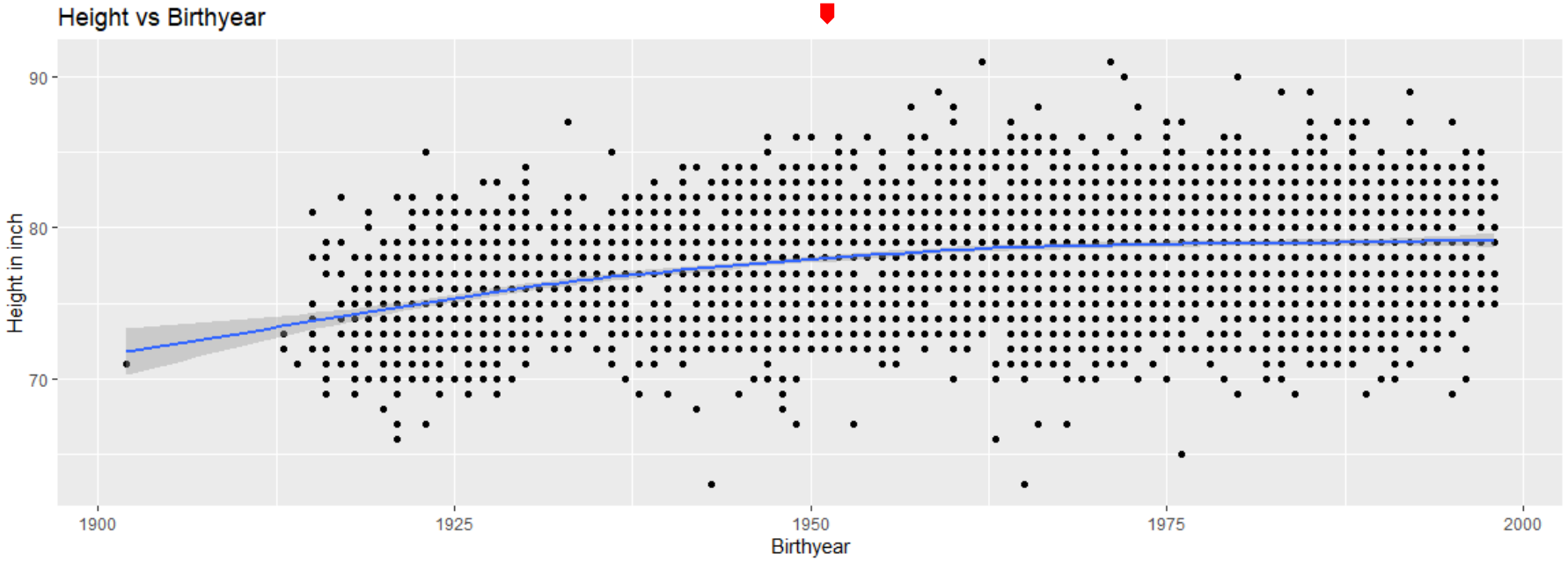
6.Height vs. Weight in Each Position

generally speaking, as height increases, weight increases

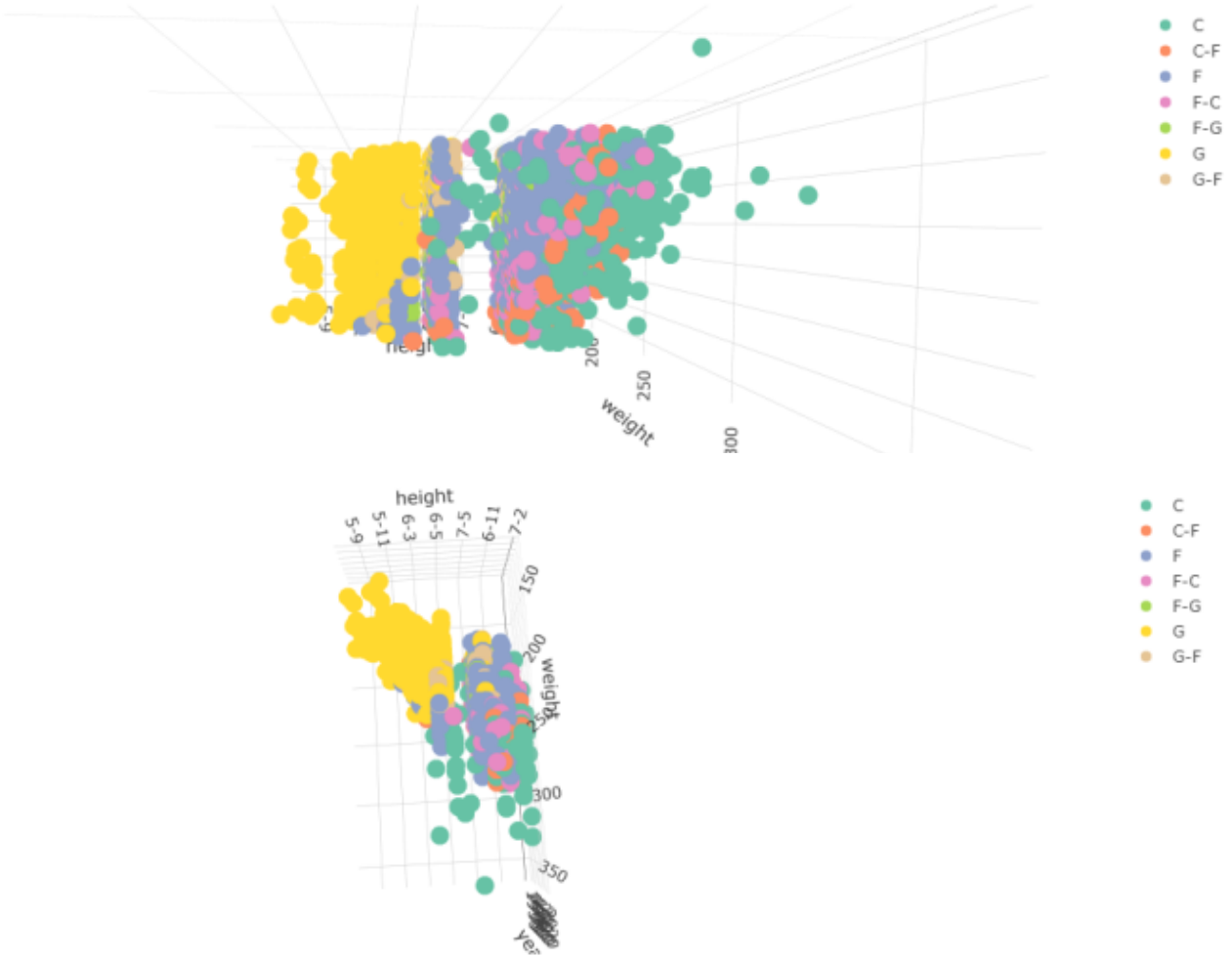


7. Height vs Years

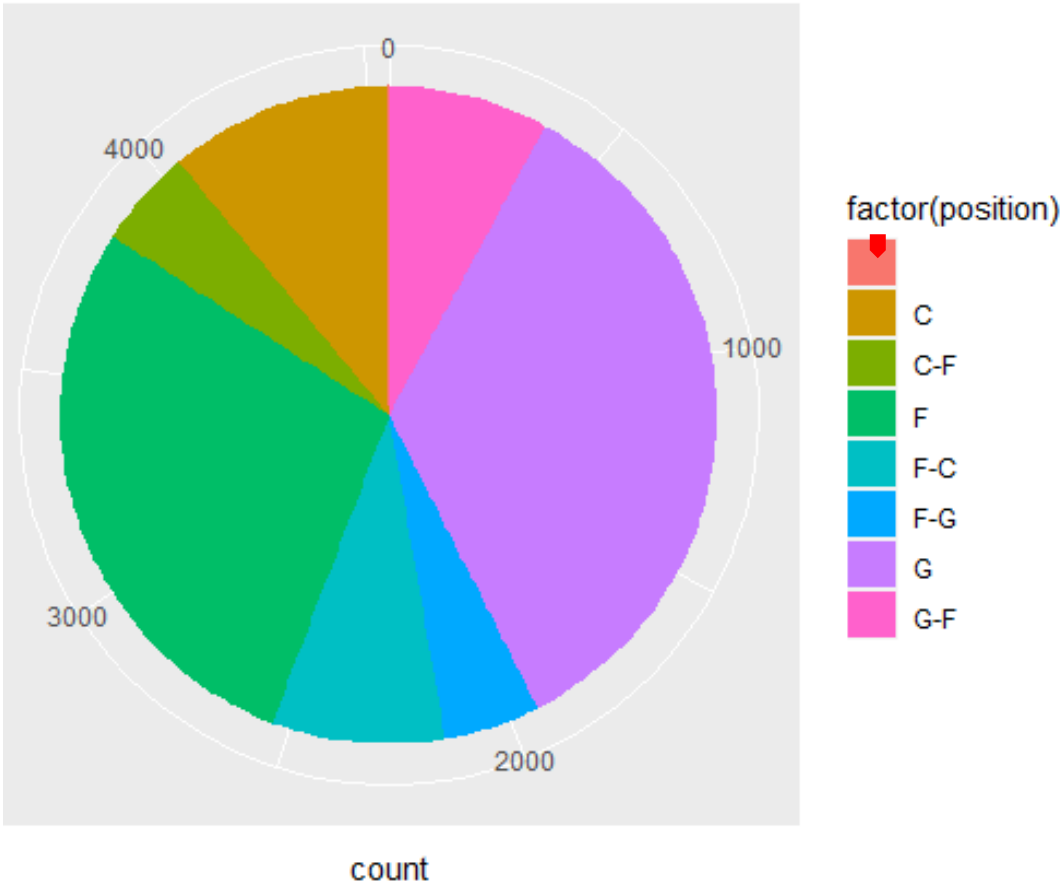
There is no visual significant difference for people born after 1950



8. 3D plot (Height vs, Weight vs, Year)



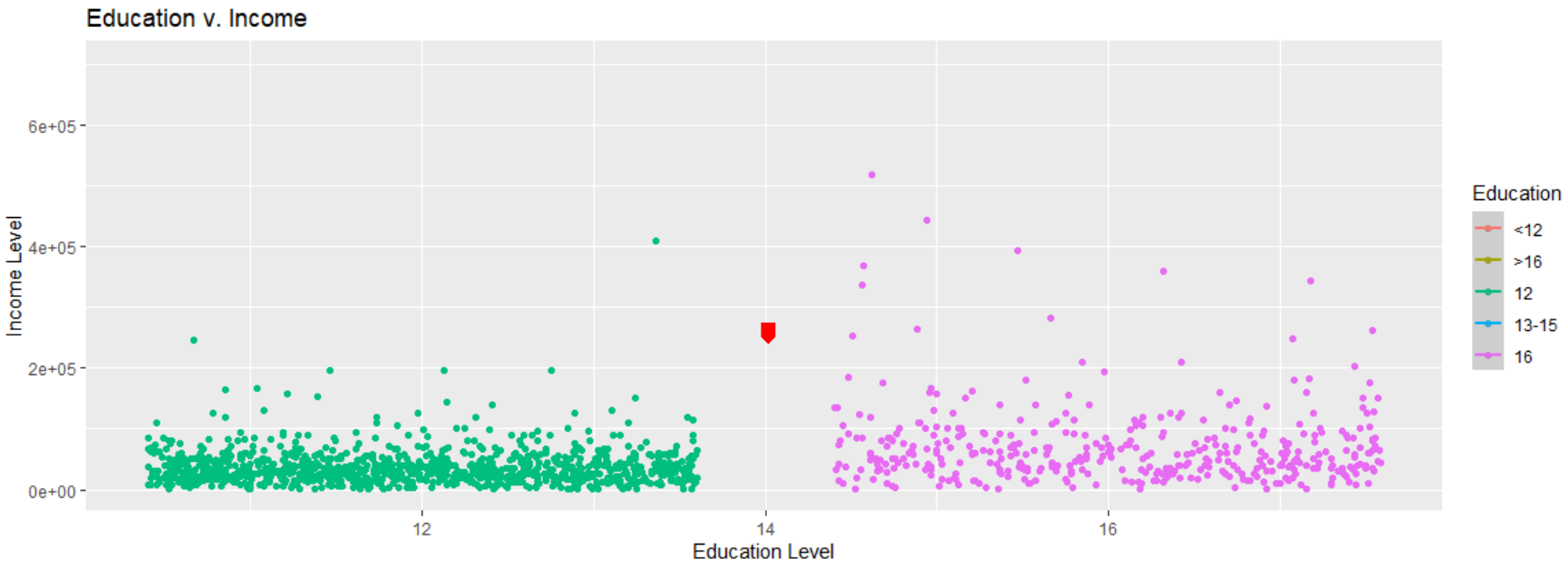
9. Using Top50 ggplot2 to present Insight



- The number of "G" and "F" are clearly the largest.
- Position "F-G" has the smallest number in the league

10.Education Level vs Income

There is no clear linear relationship between education and income (Probably incorrect)



Takeaway

For question number 2-3

How to load data specifically for position "C" and "F"

For question 10

clean column for "Education" so that it is without "<" , ">" and able to compute geom_smooth()