Thank you for applying to our Data Science position! As the next step in our interview process, we have a fun project for you to work on, to help us gauge your skills and thought pattern.

**Problem**

Can we use baseball statistics from the past 100 years to make future predictions?

**Project**

This project requires basic knowledge of the sport of baseball and statistical concepts and baseball terminology. The data set is located in this folder.

Please answer the following questions for us:

1. What are the top 3 colleges that have historically produced the best value at third base judged by their total RBI output, versus their total career salary, judged against other players at the position?
2. Is there a correlation between the year a player was born and how many walks they have in their career?
3. Based on the last 100 players who have made the all star game at shortstop, what is the most important statistical characteristic they all share and why?
4. What stadium/park do you predict will give up the most errors by a center fielder over the next 5 years and why?

**Deliverables**

Please send us the following:

1. Code, and associated files, used for the project. You can send us a zipfile, or upload the project to a public github repo.
2. Any graphs or illustrations you want to visualize the data with
3. How you analyzed the data and came to your conclusions