

# 2025 Placement Intro Project

9th June - 20th June 2025

# 1 Background

## 1.1 Introduction to Angstrom Placement

The aim of this mini project is for you to familiarise yourself with the kind of work we do at Angstrom. The following project will allow you to work on tasks that you might find yourself doing in the various teams under Modelling and Data Engineering. If you have any questions feel free to ask your buddy or your line manager.

You will have 2 weeks to complete this project, alongside shadowing someone from each of the 4 teams. On the 27th May we will ask you to do a short presentation (5-10 minutes) where you talk us through your results, thought process, methods etc. to a small internal audience. Following this, you will then roll onto your full time team and begin picking up jiras.

## 1.2 What Is the 3 Point Contest?

The 3 Point Contest is an event that occurs during the NBA All-Star Weekend.

## 1.3 Rules

Players have 70 seconds to shoot 25 balls from 5 different positions on the 3 point line + 2 extra balls (See Figure 1):



Figure 1: Shooting Locations

- On 4 of the racks, 4 of the balls are worth 1 point and 1 of the balls is a 'money ball' which is worth 2 points
- On the 5th rack, all the balls are 'money balls'. Players can choose where they want their 'money ball' rack

- There are 2 extra ‘dew’ balls which are located on either side of the top key rack. These are worth 3 points each. The name changes depending on the sponsor each year, but for the purposes of this task we will refer to them as ‘dew’ balls.
- The maximum number of points is 40.

After the qualifying round, the top 3 advance to ‘finals’. The above is repeated to get the winner.

**Extra:** in the event of a tie, multiple extra rounds of 30 seconds in the qualifying (or 1 min in finals) are played to determine the winner

## 2 Requirements

We would like you to work individually to try to predict the odds of each participating player to win the 2025 3 Point Contest. (Note that this event has already happened, but we’re going to pretend it hasn’t for the sake of this task). This task is divided into 3 subsections: Data Collection, Modelling and Simulation.

You are welcome to use whatever technologies and languages you would like.



Figure 2: 2025 Participants

## 2.1 Data Collection

Data has been provided for you, but you will have to call different API endpoints to get the data that you think could be useful for the task. You can import the Postman collection to view and call the available endpoints. Note that currently you should not be connected to any VPN in order to get the API calls to work. Things to think about include:

- Is there an efficient way of getting multiple pages of data in one go?
- How do the datasets connect with each other?
- How much of the data is relevant for what we want to do?
- Are there any outliers/data that should not be used for the modelling stage?
- What format do I want the data to be in in order to perform the modelling effectively?

If you think there is extra information you would like to use for your model or that you think is important, you are welcome to try looking for your own sources and scraping the data/calling other APIs. A suggested extra source is: Basketball Reference.

A more detailed documentation on the API is included in the project folder.

## 2.2 Modelling

This is a completely free task for you to decide how you think would be the best way to find a player's ability to shoot in the 3 point contest. There are many different ways of deciding a player's 3 make ability. You could use a rolling average or an ELO rating or something else.

Things to think about include:

- Shooting 3s in the contest is not the same as shooting 3s in a normal NBA game (e.g. there are no defenders, not all 3s shot in a real game happen exactly on the 3 point line, arm fatigue from shooting 27 balls in a row).
- In a real game, not all 3 pointers are equal in difficulty to make. Some players might have a high percentage of 3 point makes, but the majority of these might be easier, whereas other players might have a lower percentage of 3 point makes, but they go for harder shots.
- Do we just look at a player's performance in the previous season or their performance across their entire career? How would we weight the seasons if we look at their performance across multiple seasons?
- Do we take into account whether a player has participated in the 3 Point Contest in the past? If they have, do we want to take their past performance into account?

## 2.3 Simulating

Here we want to simulate the format of the contest, trying to replicate what would happen in the real world, using the model(s) that you built earlier. The output of this should allow you to reach an estimate of the probability of each player winning the contest. It should also be possible from your output to produce an estimate of the distribution of each player's possible scores.

## 2.4 Testing

If you are able to get out the probability of each player winning, you will be able to get betting odds. However, it should be noted that odds found on other bookies' websites will include margin.

$$ImpliedProbability = \frac{1}{DecimalOdds}$$

$$\frac{1}{(DecimalOdds_{optionA} * 100)} + (\frac{1}{DecimalOdds_{optionB} * 100)} = Margin$$

Using odds from other bookies (and/or the actual results of the 3 Point Contest) you should be able to try to refine your model to give better results.

## 2.5 Extension

If you have time, can you modify your work such that the following questions can be answered:

- How likely is the contest to include at least 1 player who makes both dew balls?
- How likely is each player to make all of their money balls?
- How likely is each player to advance from the qualifying round?
- How likely is each player to score 25+ points in the qualifying round?
- How likely is it that the final round record score of 29 will be broken?