Data Specific Questions

Data:

LinkedIn's team hires data scientists and researchers who work with huge amounts of data, solve real problems

1. What is the max amount of data do you handle in your role? (they should be handling large amounts of data, in petabytes: this is to understand building our large scale infrastructures)

2. What is your preference for a role- building "customer facing" applications which can be used by over 400 million members on our website or do you want to build middleware/ infrastructure for other application developers to build applications upon? In this case your customers will be internal customers. (first half is apps focused; second half is SI)

Data Phone Screen Interview (TPS)

2 questions that have to be asked: data mining and coding

CA1 (Coding algorithm): Extensive white board coding with moderate algorithmic complexity to evaluate a data candidate they mostly ask questions based on data structures, hash tables (25 mins)

Data Mining (DC): Translating data mining ideas into code and have to write at least 10 lines of code (25 mins)

*Send an email to your interviewer before the interview so they know there is a new script for the TPS and both questions have to be asked

Data Mining Questions: (under 20 mins)

-Binary classifiers

-Clustering methods

-Linear regression (supervised learning, SVM)

-Ranking methods

-Logistic Regression

-Regularization/overfitting

-Cross validation

-PCA