

**Netflix recommender project call**  
**21<sup>st</sup> June 2021**  
**Present: Michael Hoang (MH), Vivek Kotecha (VK)**

**Progress**

Not heard from RS

**Discussion on blogs**

General feeling is to have less steps in blogs as overwhelming.

Put actual steps into documentation or into guided project on Github, or link to Dataquest missions that are relevant. Videos are good for nuance like discussing analytical challenges, like assigning countries for titles with label: 'Spanish language shows'

Agreed – blog not best place to put all steps in. So use video or text to explain.

E.g. video explaining missingness. Then separate out R and Python explanations.

Video about contextualising data. What's relevant for the research problem?

**Key concepts:**

Understanding the data

Explain dataset

Missing data

Which data is relevant for our Qs

Using other data to fill missing gaps

Reading descriptions, variable types

Know your data – e.g. spot that year is wrong, for multi-season show!

Processes

Exploratory analysis

**Ideas for Tier 2**

Plotly interactive charts

Maybe a heatmap of correlation between genre and rating

**Next steps**

Follow up Rucha on progress on Tier 1

MH: edit down first 2 articles

VK: Planning videos for Tier 1 (2-3 minutes long)

**Next week to discuss**

Drafted up articles from MH. Review and make suggestions.

Put out articles in line with DQ email

Prepare python article/documentation