synthesis

We Do Human Centred Data Science

Synthesis x Data Scientist Recruitment Challenge

Thank you for your interest in Synthesis. As part of our recruitment process we set a challenge designed to test your thinking, technical ability and presentation skills. We are looking for candidates who are able to extract interesting observations from data, then give a clearly reasoned opinion about 'why' and 'what' these observations mean.

The challenge has two phases:

1. **Pre-work:** presentation + technical appendix (code/ algorithms/ pseudocode) submitted by email the morning of your interview.
    - Use any programming language/ tools you believe are suitable - be ready to describe them
    - Feel free to use code from online sources - ensure that you cite and reference appropriately
    - We encourage using data visualisations to best illustrate your findings and help you tell a compelling story around your findings/ hypotheses

2. **Interview:** present your technical challenge (<15mins) to a mixture of Data Scientist and Cultural Strategist. This can be in any format you decide. Plan to cover the following:
    - Approach: analytical/ coding techniques used, and the reasoning that drives your approach
    - Findings: what is interesting from a data and/ or a human behavioural perspective
    - Implications: what is the implication of these findings for the client and how do you foresee them to act on it?
    - Constraints and improvements: what could be done better/ differently?

We are happy to answer questions and share feedback as you work on your challenge. We encourage you to drop us a mail and take the chance to reach out for advice as you develop your approach.

Good luck!

Challenge 1: What protective facemasks are best?

Your client is a manufacturer of personal care products.

As mask wearing becomes compulsory in many areas around the world, they are looking to increase their production and sales of face masks. To inform their marketing strategy and research and development department, they want to understand the competition:

- Which are the most popular face masks out there?
- What do consumers like about them? Why?
- What different profiles of consumers buy masks? (see point on additional data)

Datasets: your colleagues have provided you with two datasets:

1. A list of facemasks on online retailer iHerb with data around the product and its price, number of ratings
2. A list of consumer reviews of these face masks

Please think of at least one **additional dataset** and provide a Proof of Concept about how it could be used to gain additional insights around the type of consumers buying different face masks.

Output: presentation + technical appendix addressing the questions above. Please explain your approach, analysis method, key findings and client implications.