Role of the Data Scientist

In Early Phase Startups

1. User Frictions
   - Understand the problem that the user needs solved
   - Understand the high-level context that will be represented by the data
   - Suggest interview questions to emerge data-related needs and current state

2. Ideation
   - Come up with ideas that could solve user frictions
   - Refine other’s ideas and product visions
   - Bring up ethical AI considerations

3. Validation
   - Understand what kinds of data may be useful to solve the use case, and which of these exist/do not exist
   - Understand where the data comes from (data inputs, IoT devices, etc) and what it means
   - Explore whether there is any data that can be purchased for said use-case
   - Explore whether there are any vendors that can already solve the data science problem at hand
   - Determine whether a data gathering campaign will be needed and how to run it
   - Assess ethical AI risks and guardrails

4. Incubation
   - Provide input for the design of the data gathering application(s) relevant
   - Analyse and understand the data and the relationship to the user need
   - Implement lean data science solution – sometimes a basic business rule is just enough to test with users
   - If a model is needed for MVP, implement a first version
   - Prepare your minimum viable software infrastructure setup
   - Bring your model to production to be used by the product, it does nobody any good sitting on your local machine
   - Iterate, pivot based on what the team learns from user testing
   - Implement ethical risk mitigations

5. Commercialisation
   - Analysis to understand how users are using the product
   - Start moving towards data engineering and pipelining for scale (where needed)
   - Start working on more advanced modelling/business rules (where needed)
   - Perform and analyse results of A/B testing
   - To implement new/alternative features