### \*\*Introduction (5 minutes)\*\*

\*You\*: "Hey everyone! How are we feeling today? Excited? Terrified? Somewhere in between?"

\*(Pause for reactions, maybe a laugh or two.)\*

\*You\*: "Don’t worry, I will try my best to not to bored you. Think of this session like a treasure hunt, except instead of gold coins, we’re hunting for data insights. We’re going to walk through this case study together, step by step."

---

### \*\*Step 1: Understanding the Problem Statement (5 minutes)\*\*

\*You\*: "Alright, step one: \*\*what’s the problem?\*\* This is like that moment when you realize your WiFi isn’t working, and you need to figure out what’s wrong before you start pressing buttons. So, let’s look at the problem."

\*\*Problem Summary\*\*:

"We’ve got a huge amount of data from the global oil and gas industry—about 80,000 assets across 3,500 companies. The big question is: what can we do with all this information? We need to analyze the data and figure out things like how much oil is being produced, which oil fields are the most valuable, and which companies own the biggest shares. The goal is to help the oil and gas companies make better decisions about where to invest, what to buy, and how to improve their performance."

\*You\*: "Basically, the big oil companies are like, 'Hey, we’ve got all this data. Can you tell us what it means?' Our job is to dig through it and figure out things. It’s like detective work, but instead of solving crimes, we’re solving oil puzzles."

---

### \*\*Step 2: Breaking the Problem Down (5 minutes)\*\*

\*You\*: "Now, imagine you’re given a huge burger—like, the biggest burger you’ve ever seen. You can’t just shove the whole thing in your mouth, right? You have to break it down into bite-sized pieces. Same thing with a case study."

\*You\*: "Let’s look at this burger, I mean, case study. Here are the \*\*bite-sized pieces\*\* we need to chew through:"

\*You\*: "So, instead of thinking, 'OMG, this is too much!' we’re just going to handle one bite at a time. Easy enough, right?"

\*(Encourage some responses, maybe add a joke about not choking on big problems.)\*

---

### \*\*Step 3: Exploring the Data (10 minutes)\*\*

\*You\*: "When we first open up the dataset, it’s kind of like that junk drawer we all have at home. There’s some good stuff in there, but also maybe a few expired coupons or random batteries. What do you think we should do first?"

\*(Wait for responses.)\*

\*You\*: "Exactly! We need to clean it up. Check for missing values, weird duplicates, and anything that doesn’t look right.

\*(Give a simple analogy of using `df.describe()` or `df.info()` as opening the fridge and checking what's edible.)\*

\*You\*: "And remember, if your data looks weird or smells funny, just like leftovers from last week, it’s probably a good idea to fix it before you start cooking!"

---

### \*\*Step 4: Designing the Approach (10 minutes)\*\*

\*You\*: "Now that we know what’s in the fridge, it’s time to come up with a plan. This is where we build the \*\*data pipeline\*\*. Don’t worry, it’s not as fancy as it sounds. It’s basically like a recipe. We’re going to lay out the steps we need to follow to get from raw ingredients (data) to a delicious dish (insights)."

\*You\*: "Here’s our recipe:"

1. \*\*Data Ingestion\*\*: This is where we grab the ingredients—aka, bring in the raw data.

2. \*\*Data Processing\*\*: Think of this as cleaning and chopping the veggies—removing the bad stuff, like missing values or duplicates.

3. \*\*Data Storage\*\*: Where do we keep the processed veggies? We organize them into different layers: raw, staging, and curated. Staging is kind of like halfway-cooked food, and curated is the final dish.

4. \*\*Data Analysis\*\*: Here’s where we do the cooking—we actually analyze the data to answer the case study questions.

5. \*\*Data Visualization\*\*: And finally, we plate the food so it looks good—this is where we make pretty graphs and charts to explain our findings.

\*You\*: "So, why do you think it’s important to follow a recipe, er, pipeline, like this?"

\*(Let them respond.)\*

\*You\*: "Right! Without a recipe, we’d be throwing ingredients into the pot at random. A pipeline keeps us organized and makes sure we don’t end up with a data soup that nobody can understand!"

---

### \*\*Step 5: Solving the Questions (15 minutes)\*\*

#### \*\*Question 1: How Much Crude Oil Will Be Produced? (5 minutes)\*\*

\*You\*: "Alright, let’s tackle the first question: how much crude oil is Denmark producing? Think of this like counting how many pizzas a restaurant has made in a year. We just need to find the right data—Denmark’s oil production—and sum it up."

\*You\*: "We’d start by filtering the data to include only Denmark. Then, we’d look at the production column and do some simple math to add it up."

\*(Explain in simple terms how filtering data works, like picking only the Denmark rows from a list.)\*

To apply the ARIMA model, the time series data needs to be **stationary** (i.e., the mean and variance should be constant over time). We can check this using the **Augmented Dickey-Fuller (ADF) Test**.

---

#### \*\*Question 2: What Are the Most Valuable Oil Fields? (5 minutes)\*\*

\*You\*: "Next up, let’s figure out which oil fields are the most valuable. This is like asking, 'Which pizza on the menu is the priciest?' We’re going to look at the financial data and see which fields bring in the most cash."

\*You\*: "We’d sort the data by revenue or valuation, then pick the top few. It’s basically like ranking pizza toppings by how much people are willing to pay for them."

\*(Explain how sorting works, like arranging pizza sizes in order of price.)\*

\*You\*: "Now, what’s tricky here? Sorting is easy, but what if some fields don’t have prices listed? Yup, just like when the pizza menu doesn’t tell you how much the large size costs. We’ll need to clean that up first!"

---

#### \*\*Question 3: Which Companies Have the Highest Ownership? (5 minutes)\*\*

\*You\*: "Finally, let’s see who owns the biggest slice of the oil pie. It’s like when you and your friends order a pizza and one friend somehow ends up with half of it. We want to see which companies have the largest ownership shares."

\*You\*: "We’d group the data by company and add up their ownership percentages. Simple, right? The company with the most slices is the winner."

\*(Explain grouping in terms of pizza slices—how many each person gets based on ownership.)\*

\*You\*: "Why does this matter? Well, the bigger the ownership, the more say they have. It’s like how the person with the most pizza slices gets to choose the toppings next time!"

---

### \*\*Step 6: Visualization and Presentation (10 minutes)\*\*

\*You\*: "Now that we’ve done all this awesome analysis, it’s time to show it off. Think of this like putting your food on Instagram—you want it to look as good as it tastes."

\*You\*: "Why do you think visualization is important?"

\*Juniors\*: "To make it easier for people to understand."

\*You\*: "Exactly! No one wants to read a giant list of numbers. We need to make it visual—like showing off those perfect pizza toppings in a pie chart."

\*(Give an example of a simple visualization using a pie chart for ownership or a line chart for production trends.)\*

\*You\*: "Different data calls for different visuals. If you’re comparing numbers, use a bar chart. If you’re showing trends over time, use a line chart. Just like you wouldn’t serve pizza on a soup plate, you need the right tool for the job."

---

### \*\*Step 7: Wrapping Up and Reflection (5 minutes)\*\*

\*You\*: "Alright, let’s do a quick recap:"

1. Understand the problem—figure out what’s broken before you start pressing buttons

.

2. Break the problem into smaller pieces—don’t try to eat the whole burger at once.

3. Explore the data—check what’s in the fridge before you start cooking.

4. Plan the approach—follow a recipe so you don’t end up with a disaster.

5. Solve the questions—filter, sort, and group your data to get answers.

6. Visualize the results—make it look good, because no one wants to see a messy plate.

\*You\*: "What’s one thing you’ll do differently next time you tackle a case study?"

\*(Let them share quick reflections.)\*

\*You\*: "Awesome! You all did great today, and remember—case studies are all about learning how to think through a problem. It’s not about rushing to the solution; it’s about breaking it down, step by step. Any final questions?"

\*(Take any final questions before wrapping up.)\*

---

### \*\*Closing (2-3 minutes)\*\*

\*You\*: "Thanks for sticking with me today, and I hope you had a bit of fun along the way. Now go forth and tackle those case studies like the data detectives you are!"

\*(End on a high note, maybe with a final joke about solving world pizza production next.)\*

---

**Analogy: The Automated Pizza Assembly Line**

Let’s say you’re running a pizza shop. You want to make sure the pizzas are being made as fast and as perfectly as possible, with minimal human error. So, you build an **automated assembly line**. Here’s how it works:

**Continuous Integration (CI) = Preparing the Pizza Ingredients**

* Imagine the **CI process** as the part of the pizza shop where you prepare and check all your ingredients.
* Every time a chef (developer) brings new ingredients (code), it’s tested right away to make sure they are fresh and meet the quality standards (the code works and doesn’t break anything).
* The dough (codebase) is continuously updated with new ingredients like cheese, toppings, and sauce (new features and bug fixes).
* If someone brings in stale or bad ingredients (bad code), the system catches it right away, and the chef is asked to fix it before moving forward.

**Continuous Deployment (CD) = Baking and Delivering the Pizza**

* Now that we have all the good ingredients, the pizza goes down the assembly line (automated process), where it gets baked and packaged automatically.
* **CD ensures** that the pizza is consistently baked to the right temperature and quality, with no manual work needed from the chefs.
* Once it’s ready, it’s delivered straight to the customer’s door (the new code goes live for users) as fast as possible.
* **The best part?** No one has to wait long for their pizza (or code) because it’s automated!
*  **Code is pushed** to a version control system (e.g., GitHub).
*  **CI tool** (e.g., Jenkins) automatically runs tests on the new code.
*  **Build automation** packages the software (e.g., using Docker).
*  **CD tool** deploys the application to a staging or production environment (e.g., using AWS CodeDeploy).
*  **Monitoring tools** (e.g., Prometheus, Grafana) ensure the app is running smoothly post-deployment.