[Skip to content](https://chatgpt.com/c/66e1ba72-df40-8007-b0e8-4bd6bdc78341#main)

**Chat history**

**You said:**

I'm a data analyst and I need to prepare documentation for a reviewer to understand how I developed my queries. I want to make sure I capture the most important items at a high level, but don't want to get bogged down in a bunch of technical details (someone can check the SQL code for that). What are the five or so most important things you think I should capture? Before you give me an answer, first ask up to three questions that will help you prepare a better response.

**ChatGPT said:**

Updated saved memory

To tailor my response better, here are a few questions:

1. What kind of data are you working with (e.g., sales, financial, healthcare)?
2. Who is the target audience for the documentation (e.g., technical peers, business stakeholders)?
3. Are there any key business objectives or questions that your queries aim to address?

These will help me focus on what’s most relevant for your situation.

**You said:**

1. Medical claims records for a three-year period. 2. My supervisor, who is a subject matter expert, but has limited data analytics skills. 3. I want my supervisor to be able to determine whether my analysis is appropriate for their needs. I don't need them to check whether I've coded things properly.

**ChatGPT said:**

Given your context, here are the top five high-level items you should capture in your documentation:

1. **Business Objective**: Clearly define the purpose of your analysis. For example, explain the specific medical claim trends or patterns you are trying to uncover over the three-year period. This will give your supervisor a clear sense of what the analysis aims to achieve.
2. **Data Sources**: Provide a summary of the datasets you used, such as claims records, patient demographics, or other relevant medical data. Mention the time period covered and any significant filtering or inclusion criteria (e.g., only patients over a certain age or specific types of claims).
3. **Key Metrics and Variables**: Highlight the primary metrics or variables you focused on (e.g., total claims cost, number of claims per patient, or types of procedures). Explain why these variables are important in addressing the business objective.
4. **Approach to Data Aggregation**: Summarize how you aggregated the data (e.g., by patient, by claim type, or by time period). This helps your supervisor understand how you've structured the data without needing to see the SQL.
5. **Assumptions and Limitations**: Outline any assumptions you've made (e.g., excluding outliers or handling missing data) and any potential limitations of your analysis that could impact the results.

These items should help your supervisor assess the relevance and soundness of your analysis without getting into technical coding details.

Top of Form

Bottom of Form

ChatGPT can make mistakes. Check important info.