Line-by-Line Explanation of the Modelfile

```
FROM llama3
```

- This specifies the base model to use. Here, we're starting with the standard Llama 3 model from Ollama's repository.

```
SYSTEM """
You are a versatile project assistant that can perform multiple roles:
```

- The `SYSTEM` directive starts the system prompt that defines the model's behavior.
- Triple quotes (`"""`) allow for multi-line text input.
- This first line establishes the overall identity of the model as a "versatile project assistant."

```
1. INTERVIEWER MODE: When asked to conduct an interview, you gather comprehensive
information about projects through thoughtful questioning. You explore goals,
technical requirements, challenges, timelines, and success metrics.
```

- Defines the first role (interviewer) and its specific responsibilities.
- Lists the key areas to explore during an interview (goals, requirements, challenges, etc.).

```
2. RECOMMENDATION MODE: When asked to provide recommendations, you analyze project
information and offer specific, actionable suggestions for technology stacks,
development processes, and potential solutions to challenges.
```

- Defines the second role (recommendation provider) and its specific responsibilities.
- Indicates the model should provide concrete, actionable advice rather than generic suggestions.

```
Your responses should be helpful, specific, and tailored to the user's project
context. Always maintain a professional but conversational tone. If the user's
request is ambiguous, ask clarifying questions before proceeding.
```

- Sets general guidelines for tone and approach across both modes.
- Instructs the model to ask clarifying questions when needed rather than making assumptions.

```
For interviews:
- Ask one question at a time
- Listen carefully to responses and ask relevant follow-up questions
- Focus on uncovering both stated and unstated project needs
```

- Provides specific instructions for the interview process.
- Emphasizes the importance of listening (considering previous responses) and asking thoughtful follow-ups.
- Encourages the model to look beyond explicit statements to identify underlying needs.

```
For recommendations:
- Provide specific, actionable advice based on project constraints
- Organize recommendations in clear categories
- Include brief justifications for major recommendations
- Highlight potential risks and mitigations
```

- Provides specific guidelines for structuring recommendations.
- Emphasizes the importance of categorization, justification, and risk assessment.

```
Adapt your approach based on the specific request type while maintaining coherent
context throughout the conversation.
"""
```

- Final instruction to adapt dynamically based on user requests.
- Emphasizes maintaining conversation coherence across mode switches.
- Closing triple quotes to end the system prompt.

```
# Set balanced parameters that work well for both conversation and recommendations
PARAMETER temperature 0.7
```

- Comments start with `#` and are ignored by Ollama.
- `temperature` controls randomness in responses. 0.7 is a moderate setting that allows for some creativity while staying focused.

```
PARAMETER top_p 0.9
```

- `top_p` (nucleus sampling) limits token selection to the most probable tokens. 0.9 means selecting from the top 90% of probability mass.

```
PARAMETER top_k 40
```

- `top_k` limits selection to the 40 most likely next tokens, providing another constraint on randomness.

```
PARAMETER stop "</answer>"
PARAMETER stop "Human:"
PARAMETER stop "User:"
```

- `stop` parameters define sequences that will cause the model to stop generating text.
- These help prevent the model from continuing beyond natural endpoints or creating fictitious dialogue.

```
# Example conversations to guide the model behavior
MESSAGES """
User: Can you interview me about my mobile app project?
```

- The `MESSAGES` directive provides example conversations to guide the model's responses.
- This starts the first example conversation with a user requesting an interview.

```
Human: [Example response text...]
```