Estimating Anthropic Fees for Coding Transcripts (December 2024)

Assumptions
-   A transcript is 100 rows (or less) of text in one Excel file
-   Each row has a sentence or thought unit or speaking turn.
-   This produces, on average, about 35,000 input tokens that are processed by Anthropic's Claude LLM. This number of tokens includes both our training instructions (which is most of the tokens) and the transcript to be coded.
-   This produces, on average, about 4000 output tokens
-   Each transcript is analyzed 5 times before a code is assigned (for an explanation, see the methods description for our models).
-   **NOTE:** If entries are longer or shorter (in terms of length of sentences, thought units, or speaking turns), costs will be <u>slightly</u> higher or lower than this estimate.
-   **NOTE:** If a transcript is more than 100 rows, we break it into two or more "chunks" (smaller files), so that the Anthropic model is never given a file longer than 100 rows. Each added file (even if it is just 10 or 15 rows) will cost about the same as a100-row file. A transcript that is 110 rows, for example, will be broken into two files of 55 rows, and will cost about twice what a transcript with 100 rows costs.

Claude 3 Opus (which is used by our Model 1)
Cost for input is $15 per 1 million tokens
Cost for output is $75 per 1 million tokens
<u>For each **10** transcripts of 100 rows</u>
- Input: 10 transcripts x 35,000 tokens x 5 (five runs per transcript) = 1.75m tokens (Cost: $26.25)
- Output: 10 transcripts x 4,000 tokens x 5 (five runs per transcript) = .2m tokens (Cost: $15)
- Total cost for 10 transcripts: $41.25

Claude 3 Sonnett (which is used by our Model 2)
Cost for input is $3 per 1 million tokens
Cost for output is $15 per 1 million tokens
<u>For each **10** transcripts of 100 rows</u>
- Input: 10 transcripts x 35,000 tokens x 5 (five runs per transcript) = 1.75m tokens (Cost: $5.25)
- Output: 10 transcripts x 4,000 tokens x 5 (five runs per transcript) = .2m tokens (Cost: $3)
- Total cost: $8.25

Examples
1. 80 transcripts of 90 rows would cost about $320 for model 1 and about $64 for model 2.
2. 80 transcripts of about 120 rows would cost about $640 for model 1 and about $128 for model 2.