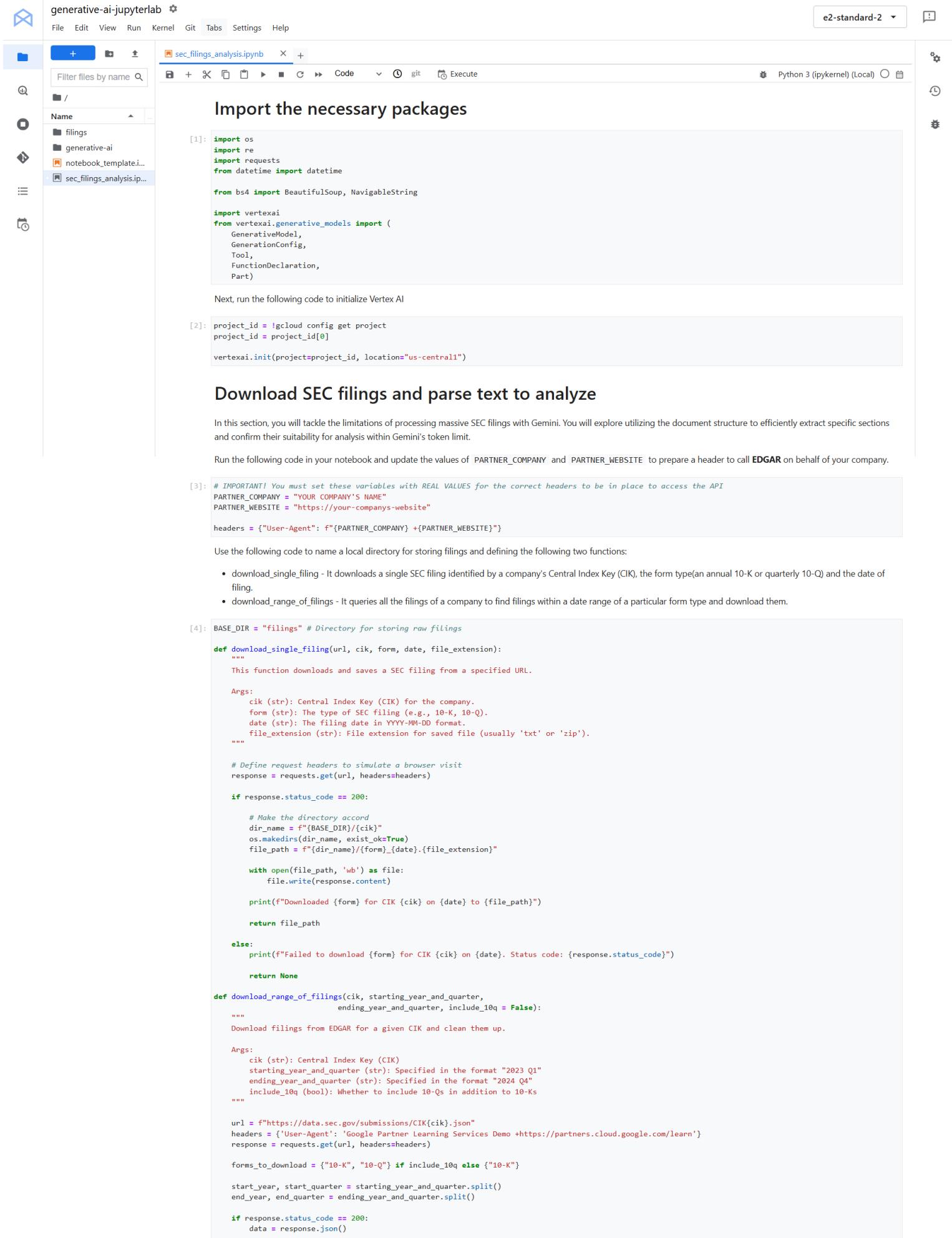


generative-ai-jupyterlab *

File Edit View Run Kernel Git Tabs Settings Help

e2-standard-2 ▾



Import the necessary packages

```
[1]: import os
import re
import requests
from datetime import datetime

from bs4 import BeautifulSoup, NavigableString

import vertexai
from vertexai.generative_models import (
    GenerativeModel,
    GenerationConfig,
    Tool,
    FunctionDeclaration,
    Part)
```

Next, run the following code to initialize Vertex AI

```
[2]: project_id = !gcloud config get project
project_id = project_id[0]

vertexai.init(project=project_id, location="us-central1")
```

Download SEC filings and parse text to analyze

In this section, you will tackle the limitations of processing massive SEC filings with Gemini. You will explore utilizing the document structure to efficiently extract specific sections and confirm their suitability for analysis within Gemini's token limit.

Run the following code in your notebook and update the values of `PARTNER_COMPANY` and `PARTNER_WEBSITE` to prepare a header to call **EDGAR** on behalf of your company.

```
[3]: # IMPORTANT! You must set these variables with REAL VALUES for the correct headers to be in place to access the API
PARTNER_COMPANY = "YOUR COMPANY'S NAME"
PARTNER_WEBSITE = "https://your-companys-website"

headers = {"User-Agent": f"{PARTNER_COMPANY} +{PARTNER_WEBSITE}"}
```

Use the following code to name a local directory for storing filings and defining the following two functions:

- `download_single_filing` - It downloads a single SEC filing identified by a company's Central Index Key (CIK), the form type(an annual 10-K or quarterly 10-Q) and the date of filing.
- `download_range_of_filings` - It queries all the filings of a company to find filings within a date range of a particular form type and download them.

```
[4]: BASE_DIR = "filings" # Directory for storing raw filings

def download_single_filing(url, cik, form, date, file_extension):
    """
    This function downloads and saves a SEC filing from a specified URL.

    Args:
        cik (str): Central Index Key (CIK) for the company.
        form (str): The type of SEC filing (e.g., 10-K, 10-Q).
        date (str): The filing date in YYYY-MM-DD format.
        file_extension (str): File extension for saved file (usually 'txt' or 'zip').
    """

    # Define request headers to simulate a browser visit
    response = requests.get(url, headers=headers)

    if response.status_code == 200:
        # Make the directory accord
        dir_name = f'{BASE_DIR}/{cik}'
        os.makedirs(dir_name, exist_ok=True)
        file_path = f'{dir_name}/{form}_{date}.{file_extension}'

        with open(file_path, 'wb') as file:
            file.write(response.content)

        print(f"Downloaded {form} for CIK {cik} on {date} to {file_path}")

        return file_path
    else:
        print(f"Failed to download {form} for CIK {cik} on {date}. Status code: {response.status_code}")

        return None

def download_range_of_filings(cik, starting_year_and_quarter,
                             ending_year_and_quarter, include_10q = False):
    """
    Download filings from EDGAR for a given CIK and clean them up.

    Args:
        cik (str): Central Index Key (CIK)
        starting_year_and_quarter (str): Specified in the format "2023 Q1"
        ending_year_and_quarter (str): Specified in the format "2024 Q4"
        include_10q (bool): Whether to include 10-Qs in addition to 10-Ks
    """

    url = f"https://data.sec.gov/submissions/CIK{cik}.json"
    headers = {'User-Agent': 'Google Partner Learning Services Demo +https://partners.cloud.google.com/learn'}
    response = requests.get(url, headers=headers)

    forms_to_download = {"10-K", "10-Q"} if include_10q else {"10-K"}

    start_year, start_quarter = starting_year_and_quarter.split()
    end_year, end_quarter = ending_year_and_quarter.split()

    if response.status_code == 200:
        data = response.json()
```

```

    filing_paths = []
    for filing_date, form, accession_number in zip(data['filings']['recent']['filingDate'], data['filings']['recent']['form'], data['filings']['recent']):
        if (form in forms_to_download) and (int(start_year) <= datetime.strptime(filing_date, '%Y-%m-%d').year <= int(end_year)):
            file_path = download_single_filing(
                f"https://www.sec.gov/Archives/edgar/data/{cik}/{accession_number.replace('-', '')}/{accession_number}.txt",
                cik, form, filing_date, 'htm'
            )
            filing_paths.append(file_path)
    return filing_paths
else:
    print("Error from call.")
    print(response)

```

Next, test the functions using the **CIK of Alphabet**, Google's parent company, which is 0001652044.

```

[5]: alphabet_cik = "0001652044"

download_range_of_filings(cik=alphabet_cik,
                          starting_year_and_quarter="2024 Q1",
                          ending_year_and_quarter="2024 Q4")

Downloaded 10-K for CIK 0001652044 on 2024-01-31 to filings/0001652044/10-K_2024-01-31.htm
[5]: ['filings/0001652044/10-K_2024-01-31.htm']

```

Open the directory `filings/0001652044`, right click on the `.htm` file and select Open in New Browser Tab to preview the document. Notice that the table of contents of these documents are highly structured and all the 10-K filings will follow the same structure

Instantiate a GenerativeModel using Gemini Pro version `gemini-2.0-flash`.

```

[6]: model = GenerativeModel("gemini-2.0-flash",
                           generation_config=GenerationConfig(temperature=0),
                           )

```

To determine if you can send the entire document to **Gemini** to analyze at once, use your GenerativeModel's `count_tokens` method to see the number of tokens in the raw file.

```

[7]: downloaded_path = "filings/0001652044/10-K_2024-01-31.htm"

with open(downloaded_path, 'r') as f:
    filing_text = f.read()

response = model.count_tokens(filing_text)
print(response)

total_tokens: 5798629
total_billable_characters: 13029439
prompt_tokens_details {
    modality: TEXT
    token_count: 5798629
}

```

With a token count of 5,798,629, you can see that even with Gemini 2.0 Flash's large token window of 1,048,576 tokens, these documents are too long to read in a single pass.

You could implement a Retrieval-Augmented Generation (RAG) framework to query small chunks of these documents, but here you are looking for a broader understanding of sections as a whole rather than smaller chunks consisting of a few facts in the document. Here, you don't mind passing a large number of tokens to Gemini, as this will be an internal tool used by a relatively small number of analysts, not a public tool handling thousands of queries.

SEC filings are required to adhere to a strict structure with named sections. You can use this standard structure of the documents to read the text between one section header and the next.

Run the following code in your notebook to define a list of the sections(items) and functions to help retrieve all of the text from one item header until the next item header.

```

[8]: items = ['Business',
            'Risk Factors',
            'Unresolved Staff Comments',
            'Properties', 'Legal Proceedings',
            'Mine Safety Disclosures',
            'Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities',
            'Management's Discussion and Analysis of Financial Condition and Results of Operations',
            'Quantitative and Qualitative Disclosures About Market Risk',
            'Financial Statements and Supplementary Data',
            'Changes in and Disagreements with Accountants on Accounting and Financial Disclosure',
            'Controls and Procedures',
            'Other Information',
            'Disclosure Regarding Foreign Jurisdictions that Prevent Inspections',
            'Directors, Executive Officers, and Corporate Governance',
            'Executive Compensation',
            'Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters',
            'Certain Relationships and Related Transactions, and Director Independence',
            'Principal Accountant Fees and Services',
            'Exhibit and Financial Statement Schedules',
            'Form 10-K Summary']

def find_div_id(soup, item_name: str):
    found_tag = soup.find('a', string=item_name)
    if found_tag:
        return found_tag["href"].strip('#')
    else:
        print(f"Couldn't find a matching tag for: {item_name}")
    return None

def get_text_between(soup, cur_name, end_name):
    cur = soup.find('div', id=find_div_id(soup, cur_name)).next_sibling
    end = soup.find('div', id=find_div_id(soup, end_name))
    while cur and cur != end:
        if isinstance(cur, NavigableString):
            text = cur.strip()
            if len(text):
                yield text
        cur = cur.next_element

def get_items_from_filings(item_names, filing_paths):
    print("Items of interest: " + ", ".join(item_names) + "\n")
    item_strings = {item: f"<{item}>\n" for item in item_names}

    for path in filing_paths:
        with open(path, 'r', encoding='utf-8') as file:
            content = file.read()

        soup = BeautifulSoup(content, 'html.parser')

        for item in item_names:
            item_index = item_strings[item]

```

```

        item_index = item_index if item_index < len(items) - 1 else 0
        item_output = ''.join(text for text in get_text_between(soup, item, items[item_index + 1]))
        item_strings[item] += f"From {os.path.basename(path)}\n" + item_output + "\n"
    return "\n\n".join(item_strings.values())

```

Next, test the above functions by running the following command and save your notebook.

```
[9]: item_names = ["Management's Discussion and Analysis of Financial Condition and Results of Operations"]
report = get_items_from_filings(item_names, [downloaded_path])

# Print the first 2,000 characters as an example.
print(report[0:2000] + "...")
```

Items of interest: Management's Discussion and Analysis of Financial Condition and Results of Operations

```
<Management's Discussion and Analysis of Financial Condition and Results of Operations>
From 10-K-2024-01-31.htm
Table of Contents Alphabet Inc. ITEM 7. MANAGEMENT'S DISCUSSION AND ANALYSIS OF FINANCIAL CONDITION AND RESULTS OF OPERATIONS Please read the following discussion and analysis of our financial condition and results of operations together with "Note about Forward-Looking Statements," Part I, Item 1 "Business," Part I, Item 1A "Risk Factors," and our consolidated financial statements and related notes included under Item 8 of this Annual Report on Form 10-K. The following section generally discusses 2023 results compared to 2022 results. Discussion of 2022 results compared to 2021 results to the extent not included in this report can be found in Item 7 of our 2022 Annual Report on Form 10-K. Understanding Alphabet's Financial Results Alphabet is a collection of businesses – the largest of which is Google. We report Google in two segments, Google Services and Google Cloud; we also report all non-Google businesses collectively as Other Bets. For additional information on our segments, see Part I, Item 1 Business and Note 15 of the Notes to Consolidated Financial Statements included in Item 8 of this Annual Report on Form 10-K. Trends in Our Business and Financial Effect The following long-term trends have contributed to the results of our consolidated operations, and we anticipate that they will continue to affect our future results: • Users' behaviors and advertising continue to shift online as the digital economy evolves. The continuing evolution of the online world has contributed to the growth of our business and our revenues since inception. We expect that this evolution will continue to benefit our business and our revenues, although at a slower pace than we have experienced historically, in particular after the outsized growth in our advertising revenues during the COVID-19 pandemic. In addition, we face increasing competition for user engagement and a...
```

Count the number of tokens in this section to see if it is under Gemini's token window of 1,048,576 tokens.

```
[10]: response = model.count_tokens(report)
print(response)

total_tokens: 12688
total_billable_characters: 50620
prompt_tokens_details {
    modality: TEXT
    token_count: 12688
}
```

Let Gemini look up companies' CIKs via Google Search

You were able to download Alphabet's annual report because you were provided its Central Index Key (CIK) number. Gemini already knows some CIKs for large public companies like Alphabet, but for some smaller companies, it may hallucinate and invent inaccurate CIKs. You can instead look up correct numbers by having Gemini use everyone's favorite well-known, public, up-to-date source of information: Google Search.

Run the following command to see Gemini provide a CIK from its internal knowledge. The CIK of Summit Therapeutics is actually 0001599298, but without the ability to look up, Gemini provides some alternative numbers.

```
[11]: model.generate_content("What is Summit Therapeutics' CIK with the SEC?").text
[11]: "Summit Therapeutics Inc.'s CIK (Central Index Key) with the SEC is **0001349558**.\n"
```

Use [Grounding with Google Search](#) to provide Gemini with a tool to conduct Google searches.

```
[12]: from google import genai
from google.genai.types import Tool, GenerateContentConfig, GoogleSearch, HttpOptions
from IPython.display import display, HTML
import os

project_id = os.environ['GOOGLE_CLOUD_PROJECT'] = 'qwiklabs-gcp-03-f435e0bdc8a5'
location = os.environ['GOOGLE_CLOUD_LOCATION'] = 'us-central1'
os.environ['GOOGLE_GENAI_USE_VERTEXAI'] = 'True'

# Initialize the client, explicitly passing project and location for Vertex AI
client = genai.Client(project=project_id, location=location, http_options=HttpOptions(api_version="v1"))
model_id = "gemini-2.0-flash"

# Configure Google Search as a tool for grounding
search_tool = Tool(
    google_search=GoogleSearch()
)

# Helper function to lookup a company's SEC CIK
def lookup_cik(company_name: str) -> str:
    prompt = f"""
Find the Securities and Exchange Commission's Central Index Key (CIK)
for {company_name}.
Return only the 10-digit integer CIK including leading zeroes.
"""

    # Generate content with grounding via Google Search
    response = client.models.generate_content(
        model=model_id,
        contents=prompt,
        config=GenerateContentConfig(
            tools=[search_tool],
            response_modalities=["TEXT"],
        )
    )

    # Concatenate all text parts of the response
    cik = ''.join(part.text for part in response.candidates[0].content.parts).strip()
    print(f"Lookup for CIK of {company_name} resulted in: {cik}\n")

    # Display the grounding metadata (search suggestions and sources)
    if hasattr(response.candidates[0], 'grounding_metadata'):
        display(
            HTML(
                response.candidates[0]
                    .grounding_metadata.search_entry_point.rendered_content
            )
        )

    return cik
```

Note: When you use Grounding with Google Search, you are required to display the corresponding Google Search Suggestions which help you understand what Google search was conducted and allow you to investigate the results. The `lookup_cik` function above includes a `display` block that renders the provided HTML in this notebook.

you can't do it without me. I'm your best bet. You're gonna need me.

Test the function and confirm you see the Google Search Suggestion results

```
[13]: lookup_cik("Summit Therapeutics")
Lookup for CIK of Summit Therapeutics resulted in: The Securities and Exchange Commission's (SEC) Central Index Key (CIK) for Summit Therapeutics is 0001599298.

 Summit Therapeutics CIK SEC Summit Therapeutics SEC filings
```

[13]: "The Securities and Exchange Commission's (SEC) Central Index Key (CIK) for Summit Therapeutics is 0001599298."

Create a FunctionDeclaration that will help Gemini know about function to lookup a CIK.

```
[14]: lookup_cik_fd = FunctionDeclaration(
    name="lookup_cik",
    description="Look up a company's CIK used for its SEC filings.",
    parameters=[
        {
            "type": "object",
            "properties": {
                "company_name": {
                    "type": "string",
                    "description": "The name of the company to look up."
                }
            },
            "required": [
                "company_name"
            ]
        }
    ],
    )
```

Note: The [search tool](#) loaded above cannot be combined with other Tools when passed to Gemini, so in order to create a Tool that combines it and other functions, you can create a dedicated model instance and invoke that via another function as you are doing here.

Empower Gemini to retrieve document sections from relevant years

In this section, you will equip Gemini to analyze specific sections of public company documents across different timeframes. You will define functions to retrieve relevant filings and extract information from desired sections within those documents for a more comprehensive analysis.

In your notebook, paste and run the following `FunctionDeclaration` that you can use to let Gemini know about the function you used earlier to download SEC filings and review its parameters. In particular, pay attention to how the `items_of_interest` section allows Gemini to provide an array of strings, but consisting only of strings matching the enum values you provide (the names of the various sections in the reports).

```
[15]: retrieve_filings_fd = FunctionDeclaration(
    name="retrieve_filings",
    description="Retrieve filings from the SEC EDGAR API for a company within a date range.",
    parameters=[
        {
            "type": "object",
            "properties": {
                "cik": {
                    "type": "string",
                    "description": "The CIK of a company whose documents will be retrieved"
                },
                "starting_year_and_quarter": {
                    "type": "string",
                    "description": "The first report quarter year and quarter in the format: 2024 Q1"
                },
                "ending_year_and_quarter": {
                    "type": "string",
                    "description": "The year and quarter in the format: 2024 Q1"
                },
                "include_quarterly_reports": {
                    "type": "boolean",
                    "description": "Whether to include 10-Q quarterly filings in addition to annual 10-K filings"
                },
                "items_of_interest": {
                    "description": "An array of one or more section (called items) of interest from 10-K or 10-Q filings",
                    "type": "array",
                    "items": [
                        {
                            "type": "string",
                            "enum": [
                                "Business",
                                "Risk Factors",
                                "Unresolved Staff Comments",
                                "Properties",
                                "Legal Proceedings",
                                "Mine Safety Disclosures",
                                "Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities",
                                "Management's Discussion and Analysis of Financial Condition and Results of Operations",
                                "Quantitative and Qualitative Disclosures About Market Risk",
                                "Financial Statements and Supplementary Data",
                                "Changes in and Disagreements with Accountants on Accounting and Financial Disclosure",
                                "Controls and Procedures",
                                "Other Information",
                                "Disclosure Regarding Foreign Jurisdictions that Prevent Inspections",
                                "Directors, Executive Officers and Corporate Governance",
                                "Executive Compensation",
                                "Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters",
                                "Certain Relationships and Related Transactions, and Director Independence",
                                "Principal Accountant Fees and Services",
                                "Exhibit and Financial Statement Schedules",
                                "Form 10-K Summary"
                            ]
                        }
                    ]
                },
                "required": [
                    "cik",
                    "starting_year_and_quarter",
                    "ending_year_and_quarter",
                    "items_of_interest"
                ]
            }
        }
    ]
}
```

Run the code snippet below to create a **Tool** that combines both of the `FunctionDeclarations` above.

Create a **system.instruction** and instantiate a new model that will follow the instructions to create analyses using the **SEC filings**. In your instructions, you will provide Gemini the current date so that it can calculate relevant dates for queries using relative terminology.

```
[17]: response1 = client.models.generate_content(
    model="gemini-2.0-flash",
    contents="Find the SEC CIK for Acme Corp. Return only the 10-digit CIK.",
    config=GenerateContentConfig(tools=[search_tool], response_modalities=["TEXT"])
)
cik = "".join(p.text for p in response1.candidates[0].content.parts).strip()

from vertexai.generative_models import GenerativeModel, GenerationConfig, Tool, FunctionDeclaration

# Define your function declarations (lookup_cik_fd, retrieve_filings_fd)
sec_tool = Tool(function_declarations=[lookup_cik_fd, retrieve_filings_fd])

system_instruction = """
- You are a research assistant to a financial analyst.
- Answer the user's question with an analysis, basing your response on information
used in filings from the SEC EDGAR database.
- Quote the SEC filing documents to support your analysis.
- If you are not certain about a CIK, use your tool to look it up.
- The current date is {current_date}.
""".format(
    current_date=datetime.today().strftime("%Y-%m-%d")
)

model = GenerativeModel(
    model_name="gemini-2.0-flash",
    generation_config=GenerationConfig(temperature=0),
    system_instruction=system_instruction,
    tools=[sec_tool]
)

response2 = model.generate_content(
    contents=f"Retrieve the 10-K for CIK {cik} covering 2023 Q4, and extract 'Risk Factors'."
)
```

When you ask Gemini a question related to a public company, it may return a response, or it may return a request for a function call. Starting with Gemini 2.0, it may ask for function calls in separate rounds of chat, or **multiple parallel function calls** in a single round of response. If multiple function calls are requested, your response **must include a function response for each function call** Gemini has requested (the Part.from_function_response section). To handle these cases, it is usually very helpful to define a handle_response function:

```
[18]: def handle_response(response):
    parts_for_inner_response = []
    for part in response.candidates[0].content.parts:
        # If the content part has an attribute called 'text'
        if hasattr(part, "text"):
            print("\n" + part.text)
        # If the content has an attribute called 'function_call'
        if part.function_call:
            function_call = part.function_call
            try:
                if function_call.name == "lookup_cik":
                    cik = lookup_cik(function_call.args["company_name"])
                    parts_for_inner_response.append(
                        Part.from_function_response(
                            name="lookup_cik",
                            response={
                                "content": cik,
                            },
                        )
                    )
                if function_call.name == "retrieve_filings":
                    context = ""
                    filing_paths = download_range_of_filings(function_call.args["cik"],
                        function_call.args["starting_year_and_quarter"],
                        function_call.args["ending_year_and_quarter"],
                        function_call.args.get("include_10q", False))
            except:
                print("No valid filings found or an error was encountered in retrieving them.")
            except AttributeError as e:
                print("Exception:")
                print(e)
            else:
                print(part)
        if parts_for_inner_response:
            inner_response = chat.send_message(parts_for_inner_response)
            handle_response(inner_response)
```

Start a chat session with the model.

```
[19]: chat = model.start_chat()
```

Run the following cell to ask a question that compares a corresponding section across two companies reports.

```
[20]: response = chat.send_message("How do Alphabet's risks in 2024 compare to Amazon's?")
handle_response(response)
```

Lookup for CIK of Alphabet resulted in: The Securities and Exchange Commission (SEC) Central Index Key (CIK) for Alphabet Inc. is 0001652044.



Lookup for CIK of Amazon resulted in: The Securities and Exchange Commission's (SEC) Central Index Key (CIK) for Amazon is 0001018724.



Downloaded 10-K for CIK 0001652044 on 2024-01-31 to filings/0001652044/10-K_2024-01-31.htm
Items of interest: Risk Factors

Downloaded 10-K for CIK 0001018724 on 2024-02-02 to filings/0001018724/10-K_2024-02-02.htm

Items of interest: Risk Factors

Based on the 2024 SEC filings, here's a comparison of the risk factors for Alphabet and Amazon:

Alphabet's Key Risks:

- * **Competition:** Alphabet faces intense competition across various industries, including online advertising and AI. They must continue to innovate to remain competitive.
 - > "We face intense competition. If we do not continue to innovate and provide products and services that are useful to users, customers, and other partners, we may not remain competitive, which could harm our business, financial condition, and operating results."^{*}
- * **Advertising Revenue Dependence:** A significant portion of Alphabet's revenue comes from advertising, making them vulnerable to reduced spending by advertisers or technologies that block ads.
 - > "We generate a significant portion of our revenues from advertising. Reduced spending by advertisers, a loss of partners, or new and existing technologies that block ads online and/or affect our ability to customize ads could harm our business."^{*}
- * **AI Investment and Ethical Concerns:** Alphabet is investing heavily in AI, which is highly competitive and rapidly evolving. They face ethical, technological, and regulatory challenges related to AI development and use.
 - > "We are expanding our investment in AI across the entire company. This includes generative AI and continuing to integrate AI capabilities into our products and services... New and evolving products and services, including those that use AI, raise ethical, technological, legal, regulatory, and other challenges, which may negatively affect our brands and demand for our products and services."^{*}
- * **Data Privacy and Security:** Concerns about data privacy and security could harm Alphabet's reputation and deter users from using their services.
 - > "Data privacy and security concerns relating to our technology and our practices could harm our reputation, cause us to incur significant liability, and deter current and potential users or customers from using our products and services."^{*}
- * **Regulatory Landscape:** Alphabet is subject to a variety of new, existing, and changing laws and regulations worldwide, particularly in data privacy, competition, and AI, which could harm their business.
 - > "We are subject to a variety of new, existing, and changing laws and regulations worldwide that could harm our business, and will likely be subject to an even broader scope of laws and regulations as we continue to expand our business."^{*}

Amazon's Key Risks:

- * **Competition:** Amazon faces intense competition across various industries and geographies, including physical, e-commerce, and cloud services.
 - > "Our businesses are rapidly evolving and intensely competitive, and we have many competitors across geographies, including cross-border competition, and in different industries, including physical, e-commerce, and omnichannel retail, e-commerce services, web and infrastructure computing services, electronic devices, digital content, advertising, grocery, and transportation and logistics services."^{*}
- * **International Operations:** Amazon's international operations expose them to risks such as local economic and political conditions, government regulations, and foreign currency exchange rate fluctuations.
 - > "Our international activities are significant to our revenues and profits, and we plan to further expand internationally. In certain international market segments, we have relatively little operating experience and may not benefit from any first-to-market advantages or otherwise succeed... Our international sales and operations are subject to a number of risks..."^{*}
- * **Variability in Retail Business:** Demand for Amazon's products and services can fluctuate significantly, impacting their operations and requiring them to manage inventory and staffing effectively.
 - > "Demand for our products and services can fluctuate significantly for many reasons, including as a result of seasonality, promotions, product launches, or unforeseeable events, such as in response to global economic conditions such as recessionary fears or rising inflation, natural or human-caused disasters (including public health crises) or extreme weather (including as a result of climate change), or geopolitical events."^{*}
- * **Fraudulent Activities of Sellers:** Amazon faces risks related to fraudulent or unlawful activities of sellers on its platform, which can harm their business and reputation.
 - > "Under our seller programs, we maintain policies and processes designed to prevent sellers from collecting payments, fraudulently or otherwise, when buyers never receive the products they ordered or when the products received are materially different from the sellers' descriptions, and to prevent sellers in our stores or through other stores from selling unlawful, counterfeit, pirated, or stolen goods... When these policies and processes are circumvented or fail to operate sufficiently, it can harm our business or damage our reputation and we could face civil or criminal liability for unlawful activities by our sellers."^{*}
- * **Data Loss and Security Breaches:** Amazon's collection, processing, and storage of large amounts of data make them vulnerable to data loss and security breaches, which could have significant consequences.
 - > "Because we collect, process, store, and transmit large amounts of data, including confidential, classified, sensitive, proprietary, and business and personal information, failure to prevent or mitigate data loss, theft, misuse, unauthorized access, or other security breaches or vulnerabilities affecting our or our vendors' or customers' technology, products, and systems, could... adversely affect our operating results; result in litigation, liability, or regulatory action... deter customers or sellers from using our stores, products, and services; and otherwise harm our business and reputation."^{*}
- * **Government Regulation:** Amazon is subject to evolving government regulations, particularly concerning the internet, e-commerce, and AI, which could harm their business.
 - > "We are subject to general business regulations and laws, as well as regulations and laws specifically governing the internet, physical, e-commerce, and omnichannel retail, digital content, web services, electronic devices, advertising, artificial intelligence technologies and services, and other products and services that we offer or sell... Unfavorable regulations, laws, decisions, or interpretations by government or regulatory authorities applying to these laws and regulations, or inquiries, investigations, or enforcement actions threatened or initiated by them, could cause us to incur substantial costs, expose us to unanticipated civil and criminal liability or penalties (including substantial monetary fines), diminish the demand for, or availability of, our products and services, increase our cost of doing business, require us to change our business practices in a manner materially adverse to our business, damage our reputation, impede our growth, or otherwise have a material effect on our operations."^{*}

In Summary:

Both companies share common risks related to competition, data security, and evolving regulations. However, Alphabet's risks are more focused on advertising revenue and the ethical considerations surrounding AI, while Amazon's risks are heavily influenced by their vast retail operations, international exposure, and the challenges of managing a large third-party seller ecosystem.

You can also compare sections across time. Run the following code snippet to see an example of Gemini comparing one section across multiple 'versions' (or filing years) of a document.

```
[21]: chat = model.start_chat()  
response = chat.send_message("How has Home Depot changed the way it describes its business over the past 3 years?")  
handle_response(response)
```

Lookup for CIK of Home Depot resulted in: The Securities and Exchange Commission's (SEC) Central Index Key (CIK) for Home Depot is 0000354950.

 Home Depot SEC CIK number

Downloaded 10-K for CIK 0000354950 on 2025-03-21 to filings/0000354950/10-K_2025-03-21.htm
Downloaded 10-K for CIK 0000354950 on 2024-03-13 to filings/0000354950/10-K_2024-03-13.htm
Downloaded 10-K for CIK 0000354950 on 2023-03-15 to filings/0000354950/10-K_2023-03-15.htm
Downloaded 10-K for CIK 0000354950 on 2022-03-23 to filings/0000354950/10-K_2022-03-23.htm
Items of interest: Business

Over the past three years, Home Depot has consistently emphasized its interconnected shopping experience and strategic investments in supply chain and technology to enhance customer experience. Here's a breakdown of the key changes:

- * **Focus on Interconnected Experience:** All filings from 2022 to 2025 emphasize the interconnected shopping experience, highlighting the integration of digital and physical channels. For instance, the 2025 10-K states, "We continued our strategic investments aimed at creating an interconnected, frictionless shopping experience that enables our customers to seamlessly blend the digital and physical worlds." This theme is consistent across the 2024 and 2023 filings as well.
- * **Emphasis on Pro Customers:** There's an increasing emphasis on serving professional customers ("Pros"). The 2025 filing mentions the acquisition of SRS, a distributor serving professional roofers, landscapers, and pool contractors, indicating a strategic move to expand services for Pros. The 2024 filing also highlights initiatives designed to drive growth with Pros, such as providing a customized online experience and a dedicated sales force.
- * **Supply Chain Investments:** All filings highlight ongoing investments in the supply chain to improve delivery speed and reliability. The 2025 filing notes, "We continue to focus on building best-in-class competitive advantages in our supply chain to be responsive to our customers' expectations for how, when and where they choose to receive our products and services." Similar statements appear in the 2024 and 2023 filings.
- * **Sustainability and Human Capital Management:** There's a consistent focus on sustainability and human capital management, organized around three pillars: Focus on Our People, Operate Sustainably, and Strengthen Our Communities. The 2025 filing states, "We view sustainability and human capital management matters through the lens of our business, with an understanding that if we support our associates, our customers, our suppliers, and the communities we serve, we also support our business and create value for our shareholders."
- * **Expansion and New Stores:** In the 2023 and 2024 filings, Home Depot mentions plans to open new stores. Specifically, the 2024 filing says, "In fiscal 2023, we initiated a plan to open approximately 80 new stores over a five-year period to address those opportunities."

In summary, Home Depot's business description has evolved to highlight its focus on an interconnected shopping experience, strategic investments in its supply chain, serving professional customers, and commitment to sustainability and human capital management. The company also emphasizes its plans for expansion with new store openings.

[:]: [REDACTED]

Simple 0 \$ 1 ⌂ ☰ Python 3 (ipykernel) (Local) | Idle Mode: Command ✘ Ln 2, Col 11 sec_filings_analysis.ipynb 3 🔍