Synthesizing question answering data from financial documents: An End-to-End Multi-Agent Approach

Chetan Harsha, Karmvir Singh Phogat, Sridhar Dasaratha, Shashishekar Ramakrishna
EY Global Delivery Services India LLP
{Chetan.Harsha,Karmvir.Phogat,Sridhar.Dasaratha,Shashishekar.R}@gds.ey.com

Abstract

Answering complex questions that require numerical reasoning over financial documents is challenging due to the diverse and scattered nature of relevant information. While large language models (LLMs) excel at financial reasoning, their enterprise deployment is often limited by cost and latency. Small language models (SLMs) present a cost-effective alternative but need to be fine-tuned with high-quality, domain-specific question-answer (QA) data. Acquiring such data requires manual expert annotation, presenting a bottleneck to the wider application of SLMs.

This work introduces a modular, scalable end-to-end agentic pipeline that extracts and selects relevant content from unstructured financial documents and then generates QA pairs from the selected content for SLM fine-tuning. Compared to the same models trained on previous manually generated data for the task, one of the models trained on our pipeline-produced synthetic data achieved competitive in-distribution performance, and all tested models demonstrated superior generalization. The framework thus demonstrates considerable potential to accelerate the deployment of smaller, cost-effective models by reducing manual data creation efforts.

1 Introduction

Answering questions requiring numerical reasoning over business documents such as annual reports, financial filings and tax documents is a challenging problem. These documents contain critical structured and unstructured information – text, tables and figures – that users often need to query for analysis or decision making. Building reliable QA systems for such documents requires domain-specific understanding, numerical reasoning, and the ability to process complex layouts.

Large language models (LLMs) have shown remarkable general reasoning ability achieving excellent performance in financial reasoning (Chen et al., 2023). However, deploying them for enterprise use-cases remains costly and often constrained by latency, privacy, and regulatory requirements. Small language models (SLMs) offer a cost-effective and customizable alternative that can be operated entirely within an organization's infrastructure, addressing many of these constraints. When data sets manually created by domain experts are available, they can be further enhanced with reasoning demonstrations generated by an LLM and then used to train the SLM (Magister et al., 2023). Using such an approach, promising results have been achieved on various tasks (Ho et al., 2023) including financial reasoning (Phogat et al., 2024). Despite these results, the requirement of a high quality manually curated data by domain experts limits the wider applicability of SLMs for specific domains and tasks in industry.

Synthetically generating entire training data sets is an attractive approach (Wang et al., 2023; Ye et al., 2022; Wang et al., 2021; Gou et al., 2021), but achieving quality data remains challenging (Gandhi et al., 2024) and current methods yield mixed results (Ding et al., 2023). Creating synthetic data from business documents is particularly complex due to the relevant information being scattered across text and tables in documents with complex layouts. Building diverse, meaningful numerical reasoning QA pairs from financial documents usually requires substantial manual effort, as seen in datasets like FinQA (Chen et al., 2021) and TAT-QA (Zhu et al., 2021), where financial experts spend considerable time crafting the multi-step reasoning questions. While the process yields high-quality data, it is time consuming, costly and hard to scale across different business domains.

In this work we develop and evaluate an integrated end-to-end pipeline that generates fine-tuning data from unstructured financial documents. The pipeline begins with a large corpus of business

Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 5: Industry Track), pages 669–687

documents, extracts text and tabular content and employs an LLM -based agent to identify and select the most relevant content. Additional agents then generate question-answer pairs from the selected content, while human review and feedback can be incorporated in any stage. This design enables scalable, iterative dataset creation with significantly reduced manual effort.

Our contributions are summarized below:

- **Agentic workflow:** We present a modular, multi-agent workflow that transforms financial documents into question-answer pairs suitable for training models for numerical reasoning.
- **Evaluation on realistic financial QA tasks:** We study the practical effectiveness of our framework by applying it to a large corpus of financial documents and using the resulting data to train three different SLMs. The performance is compared to that of the same SLMs trained on previously existing manually curated data.
- **Empirical insights:** We show that, for in-distribution data, one of the SLMs ($\sim$4 billion parameters.) trained with data generated by our pipeline achieves comparable performance to the same model trained on real data. Moreover, all SLMs trained using pipeline-generated data showed notable performance gains in cross-data set generalization over the models trained on real data, suggesting that our pipeline offers an efficient solution for scalable, high-quality data generation in domain-specific financial reasoning tasks.

2 Related Work

LLMs have been used to generate synthetic data or augment existing data sets for various problems such as text classification (Li et al., 2023; Yu et al., 2024), mathematical reasoning (Luo et al., 2023) and question answering (Li and Tajbakhsh, 2023; Wu et al., 2024; Schmidt et al., 2024). None of these studies focus on synthetic data generation for numerical reasoning over financial reports.

Recent studies generate synthetic financial QA data using financial formulas (Yuan et al., 2024), generate new contexts for questions in an existing data set (Hwang et al., 2023) or use passages from existing data sets to generate QA pairs (Harsha et al., 2025). Unlike these approaches we do not rely on external financial knowledge, or preexisting data sets and generate the QA data directly from raw financial documents.

LLM-based multi-agent frameworks have been introduced to improve synthetic corpora (Abdullin et al., 2023; Ge et al., 2025; Ye et al., 2025; Mitra et al., 2024). (Liu et al., 2025) note that existing datasets focus on general instruction data without domain-specific constraints, and introduce AgenticMath—a pipeline for creating high-quality math question-answer pairs to better fine-tune LLMs. However, none of these works study the generation of financial QA data, nor do they address end-to-end data creation from business documents.

(Miao et al., 2025) introduce Easy Dataset, an LLM-based framework for generating QA pairs from unstructured documents. Their approach is domain-agnostic and relies on persona-driven prompting, rather than addressing the specific requirements of numerical reasoning in the financial domain.

In summary, while previous studies have explored aspects of synthetic data generation for financial QA, we develop an end-to-end agentic pipeline and benchmark the performance of SLMs trained on generated data against those trained on real datasets, yielding important practical insights.

3 Methodology

The proposed framework (Figure 1) comprises a set of specialized agents responsible for content extraction, content selection, question and answer synthesis, and validation.

We separate the extraction and selection agents because they serve distinct purposes and have different computational requirements: extraction is compute-intensive, whereas selection is lightweight and can be scaled independently. Question and answer generation are likewise decoupled from validation to promote diversity during generation and correctness during verification. The modular design allows independent optimization and targeted refinements without rerunning the full pipeline. All agents use GPT-4o, with the extraction agent leveraging the multimodal capabilities. The framework can be easily extended to include human feedback during question and answer generation.

3.1 Content Extraction Agent

The Content Extraction Agent extracts relevant text and tables from unstructured financial reports by

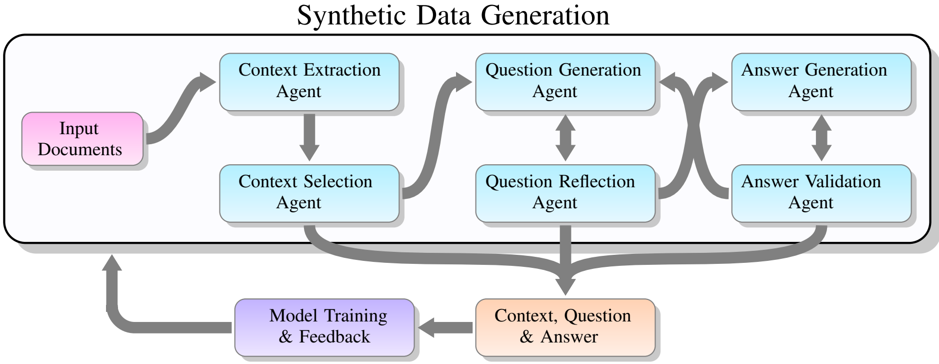

Figure 1: Workflow for synthetic data generation and SLM finetuning for financial question answering.

processing semantically coherent units—like sections, paragraphs, or smaller chunks—based on document layout. Leveraging GPT-4o 's multimodal abilities, it retrieves both text and tables from page images, standardizing table formats. Each segment may be tagged with certain attributes, aiding later processing; tagging at this stage benefits from layout cues interpreted by the multimodal LLM. The complete prompt is provided in Figure 2, Appendix A.

3.2 Content Selection Agent

The Content Selection Agent plays a critical role in ensuring QA data quality by assessing whether each extracted segment is suitable for producing multi-hop numerical reasoning questions. Without it, pages unsuitable for generating appropriate financial reasoning QA pairs could be utilized, lowering the data quality. After this step, the pages identified as irrelevant are filtered out and only the information-rich segments are sent for further processing. The prompt is provided in Figure 3 in Appendix A.

3.3 Question Generation Agent

The Question Generation Agent produces financial reasoning questions from the selected segments. Using LLMs guided by custom prompts, the agent generates questions targeting boolean or numerical answers that require multi-step reasoning across text and tables. We adapt the prompt from (Harsha et al., 2025) to include the feedback from the reflection agent (Figure 4 in Appendix A). The reflection agent evaluates each question for clarity, answerability, and consistency, refining or regenerating low quality questions. The prompt used by reflection agent is provided in Figure 5 in Appendix A.

3.4 Answer Generation Agent

The Answer Generation Agent synthesizes executable Python code for each generated financial question. Given the question and its associated context, the agent generates code that performs the arithmetic or logical operations needed to compute the final answer. The prompt is adapted for answer generation from (Harsha et al., 2025).

3.5 Answer Validation Agent

The Answer Validation Agent checks each generated QA pair by running the related Python code and ensuring it produces a valid scalar output. Code that doesn't execute, causes errors, or returns non-scalar outputs (like lists or dictionaries) is rejected. If validation fails, the system triggers targeted regeneration: the answer is regenerated to fix calculation errors, or the question is revised if its formulation is flawed. This process repeats for a set number of attempts; unresolved cases are then discarded. Only executable, well-formed, and financially relevant examples are kept for model training.

3.6 Model Training and Feedback

The final curated dataset is used to fine-tune SLMs, with evaluation performed on a held-out subset of samples. If performance thresholds are not met, earlier components of the pipeline can be refined by expanding data coverage, adjusting prompts, or modifying validation criteria. This feedback-driven loop supports continuous improvement of data quality and model performance.

4 Experiments

We evaluate the proposed framework by generating synthetic financial question-answering datasets and benchmarking them against the manually curated FinQA dataset (Chen et al., 2021), which contains expert-annotated real-world examples. The pipeline operated without any human review or feedback. SLMs were fine-tuned on both datasets and tested on the real FinQA test set to compare performance. Additionally, generalization was assessed using the independent TATQA (Zhu et al., 2021) test data.

We begin with FinTabNet (Zheng et al., 2021), a large corpus of earnings reports used to create FinQA. The FinTabNet corpus comprises 89,946 pages extracted from earnings reports of 428 S&P 500 companies (1999-2019). To prevent data leakage and ensure a fair evaluation, we exclude all 2,110 pages that were used in the construction of the FinQA dataset. This precaution is necessary because large language models may have been exposed to these passages during prior training, which could make it easier for them to generate questions similar to those in FinQA. After this exclusion, 87,836 pages remain in our corpus. From this corpus, we randomly sampled documents while ensuring representation across all companies to preserve diversity in reporting formats and content styles.

To remain consistent with the FinQA dataset construction process, we fix the extraction unit to a single page and include instructions in the prompt to tag the page as simple or complex. The complexity criteria mirror those given to FinQA annotators. A page is labelled complex if it contains multiple tables, a table with more than twenty rows, nested or hierarchical structures, or catalog-style content. All remaining pages are treated as simple and used for downstream question–answer generation to match the FinQA's selection methodology. Three questions are generated per page, and subsequent agents process them as described in the methodology to produce the final training dataset.

Using this synthetic dataset, we fine-tuned three SLMs: PHI-3-MINI, SMOLLM-1.7B, and SMOLLM-360M. For comparison, we fine-tuned the same models with the official FinQA training data set. Fine-tuning followed the setup and hyperparameters selection approach of (Phogat et al., 2024). For the SMOLLM models we chose learning rate as 3e-4 as per (Allal et al., 2025). Training was conducted for four epochs using the vLLM framework on a system with 24 CPU cores, 220 GB RAM, and a single A100 GPU (80 GB).

During evaluation, for each question in the official test sets of the FinQA and TATQA benchmarks, the fine-tuned models generated Python code to compute an answer, which was then executed and compared against the ground-truth values.

5 Results

5.1 Extraction and Chunk selection

From 87,836 available pages, we randomly sampled 12,000 pages across 428 companies and multiple years to increase diversity. The Content Extraction Agent successfully processed 92.4% (11,088 pages), with about 900 pages failing due to API errors or filtering constraints. Using provided classification criteria, 3,276 pages were labelled as simple and 7,360 as complex. Of the simple pages, 19% (621) were invalid for financial question generation, resulting in 2,655 valid simple pages; only 60 complex pages were filtered out as invalid. See Appendix B for representative examples.

To evaluate agent performance, 50 samples were manually reviewed: invalid simple pages were typically tables of contents, index pages, or lacked financial data; invalid complex pages were usually large tables without reasoning data or contained descriptive details of executives. Only one misclassification occurred, showing the agents applied filtering and classification constraints with high reliability.

5.2 Evaluation of generated data

To maintain consistency with prior research by (Harsha et al., 2025), for training with real data we use 5698 data points from the official FinQA training data. We randomly selected 5,698 samples from the pool of generated synthetic data, matching the number of real FinQA data points used.

Table 1 summarizes the performance of the three SLMs trained with the synthetic data as well as the real FinQA training data set. The performance of the models is measured on the official FinQA and TATQA test data sets.

The PHI-3-MINI model trained on the synthetic data achieves an accuracy of 64.9% on the real FinQA test data, which is within 4% of that achieved with the real data. These results demonstrate the effectiveness of the multi-agent data generation pipeline in generating fine-tuning data for financial question answering. Interestingly for the TATQA data set, the model based on the synthetic data achieves 3% higher accuracy than the model trained with real data, hinting at better performance on out-of-distribution data.

For the FinQA test data set, the performance gap between the synthetic data and real data trained

Finetuned SLMs	Training Dataset: Synthetic FinQA*		Training Dataset: Real FinQA	
	FinQA Test Accuracy	TATQA Test Accuracy	FinQA Test Accuracy	TATQA Test Accuracy
PHI-3-MINI	64.9	83.5	68.4	80.1
SMOLLM-1.7B	44.9	69.2	55.8	60.0
SMOLLM-360M	16.3	31.6	27.5	20.0

* The training datasets — real and synthetic FinQA — each contain 5,698 samples.

Table 1: Comparison of models trained on synthetic and real data for financial question answering.

Finetuned SLMs	FinQA Test Accuracy			TATQA Test Accuracy		
	Syn FinQA: 10k*	Syn FinQA: 20k*	Real	Syn FinQA: 10k*	Syn FinQA: 20k*	Real
PHI-3-MINI	66.2	65.2	68.4	83.7	83.9	80.1
SMOLLM-1.7B	47.2	46.5	55.8	74.1	75.1	60.0

* The training datasets include two synthetic FinQA versions — Syn FinQA: 10k and Syn FinQA: 20k, containing 10,000 and 20,000 samples respectively — along with the real FinQA dataset, which contains 5,698 samples.

Table 2: Comparison of models trained on synthetic (various sizes) and real data for financial question answering.

models is higher for the smaller models, as compared to the PHI-3-MINI model. The SMOLLM models trained on synthetic data fall short by more than 10% of the accuracy achieved with the real data. The performance on the TATQA test data set again reveals notably better performance with the synthetic data trained SMOLLM models surpassing the corresponding models based on real data by more than 9%. This further provides evidence of the better out of distribution performance achieved by the synthetic data trained models.

5.3 Effect of sample size

To assess whether increasing synthetic data enhances model performance, we fine-tuned PHI-3-MINI and SMOLLM-1.7B models using larger synthetic datasets of 10,000 and 20,000 samples, compared to the baseline of 5,698 samples. Table 2 shows the results for the FinQA and TATQA data sets. On the FinQA test set, expanding sample size yielded slight improvement for both models. However, for the TATQA test set, the SMOLLM-1.7B model benefited substantially from more data. When trained with 20,000 samples it achieved a nearly 6% gain over the baseline and 15% higher accuracy than the model trained on real data. Its performance also came within 5% of the much larger PHI-3-MINI model trained on real FinQA data. These findings show that increasing synthetic data generated by our framework enables smaller models to perform competitively, narrowing the gap with larger models.

5.4 Performance analysis

We further evaluated system performance by comparing accuracy across various dimensions. The FinQA datasets categorize questions based on the source of required information: table-only, text-only, and text-table questions. Table 3 shows that the performance gap between models trained on synthetic versus real data is largest for table-only questions, while for text-only questions, the gap is smaller, and two models even perform as well or better with synthetic data. We also analyzed results by question complexity as measured by the number of steps required to answer the question: 1-step, 2-step, and more than 2-step questions (Table 4). SMOLLM models trained on synthetic data lag in all complexity categories. The PHI-3-MINI model displays similar performance for higher-complexity questions, but synthetic data slightly underperforms for the 1-step questions. One possible reason for these observations could be that the FinQA data is skewed towards single and two step questions while our prompt asks for multi-hop question generation without enforcing any constraints in terms of number of steps required to solve the problem.

5.5 Discussion

The original FinQA dataset was created through an intensive process in which professional financial experts with CPA or MBA backgrounds, generated QA pairs, after receiving specialized training. The data collection period lasted about eight weeks. In addition, the source documents in the FinTabNet

Finetuned SLMs	Original FinQA			Synthetic FinQA		
	Table	Text	Table & Text	Table	Text	Table & Text
PHI-3-MINI	75.0	60.2	53.7	69.4	60.8	51.9
SMOLLM-1.7B	62.0	47.7	43.0	48.7	42.0	32.0
SMOLLM-360M	34.9	16.9	13.2	17.1	19.7	6.3

Table 3: Comparison of FinQA test accuracy across different data modalities (Table, Text, and Table & Text) for original and synthetic FinQA datasets.

Finetuned SLMs	Original FinQA			Synthetic FinQA		
	1 Step	2 Steps	>2 Steps	1 Step	2 Steps	>2 Steps
PHI-3-MINI	72.1	67.2	45.2	67.2	64.8	46.4
SMOLLM-1.7B	55.8	58.1	45.2	44.3	48.6	30.9
SMOLLM-360M	29.6	27.8	9.0	15.7	18.3	6.0

Table 4: Comparison of FinQA test accuracy across reasoning step complexity (1, 2, and >2 steps) for original and synthetic FinQA datasets.

data were annotated which was used for filtering pages during the creation of FinQA dataset. The thorough human annotation resulted in high data quality but demanded significant cost and domain expertise.

In contrast, we do not use any manual annotation, instead relying on autonomous agents for parsing documents, selecting content, and generating QA data. Instead of using FinTabNet annotations, the Content Extraction and Selection Agents jointly handle filtering to select pertinent report pages, while other agents create reasoning-driven QA pairs. Human input is limited to prompt design and output evaluation, markedly reducing manual effort. Moreover, this approach can be easily adapted to other use-cases within finance or similar tasks in other domains by adjusting agent prompts as needed.

Our pipeline incorporates feedback loops after question as well as answer generation, which improve data quality. However, similar quality can be achieved by generating more samples than necessary and filtering out those rejected by the Answer Validation Agent to reach the target number of quality samples. When the document corpus is much larger relative to the sample requirement, this simpler approach is just as effective. Conversely, with a smaller document set, feedback loops help maximize use of available raw data.

Unlike FinQA, the TATQA test set was compiled from a separate document corpus, annotated by a different team using distinct criteria for page selection. Notably, our approach yielded considerable performance gains on the TATQA set, especially for the two smallest models. This is a key finding given that our models relied on the same raw data sources and instructions in the prompts aligned closely with those used for FinQA annotation. These findings indicate that synthetic data generated by our framework can improve model generalizability and applicability compared to models trained solely on real data.

6 Conclusion

We developed an integrated, end-to-end agentic pipeline capable of generating fine-tuning QA datasets directly from unstructured financial documents, substantially reducing the need for manual data collection and annotation. Our experiments with three SLMs trained on synthetic data produced by this pipeline demonstrated that one model achieved in-distribution performance close to the same model trained on expert-annotated datasets, and, importantly, all three models exhibited superior generalization capabilities compared to their counterparts trained on real data. Notably, increasing the volume of generated data led to substantial performance improvements in generalization for the smallest models, indicating that our approach supports the practical deployment of smaller, more cost-efficient models for specialized financial tasks.

Limitations

All agents in our pipeline use GPT-4O. This means the pipeline inherits the model's strengths and blind spots, which may limit coverage of certain financial questions and associated reasoning patterns.

Using diverse LLMs and including human-in-the-loop can alleviate some of these concerns. While the final SLMs can be run on-prem, the agentic pipeline uses a cloud based LLM. As we used publicly available financial reports this was not a concern but other use-cases with strict data restrictions, may need the use of on-prem open source LLMs to power the agents, necessitating further studies to measure their effectiveness. We evaluated the performance of the current system for numerical reasoning over financial documents. Hence our empirical findings while highly encouraging must be limited to the specific domain considered, while motivating application and research in other domains.

Disclaimer

The views reflected in this article are the views of the authors and do not necessarily reflect the views of the global EY organization or its member firms.

References

Yelaman Abdullin, Diego Molla, Bahadorreza Ofoghi, John Yearwood, and Qingyang Li. 2023. Synthetic dialogue dataset generation using LLM agents. In *Proceedings of the Third Workshop on Natural Language Generation, Evaluation, and Metrics (GEM)*, pages 181–191, Singapore. Association for Computational Linguistics.

Loubna Ben Allal, Anton Lozhkov, Elie Bakouch, Gabriel Martín Blázquez, Guilherme Penedo, Lewis Tunstall, Andrés Marafioti, Hynek Kydlíček, Agustín Piqueres Lajarín, Vaibhav Srivastav, Joshua Lochner, Caleb Fahlgren, Xuan-Son Nguyen, Clémentine Fourrier, Ben Burtenshaw, Hugo Larcher, Haojun Zhao, Cyril Zakka, Mathieu Morlon, Colin Raffel, Leandro von Werra, and Thomas Wolf. 2025. SmolLM2: When Smol Goes Big – Data-Centric Training of a Small Language Model. *arXiv preprint arXiv:2502.02737*.

Wenhu Chen, Xueguang Ma, Xinyi Wang, and William W. Cohen. 2023. Program of Thoughts Prompting: Disentangling Computation from Reasoning for Numerical Reasoning Tasks. *Transactions on Machine Learning Research*.

Zhiyu Chen, Wenhu Chen, Charese Smiley, Sameena Shah, Iana Borova, Dylan Langdon, Reema Moussa, Matt Beane, Ting Hao Huang, Bryan Routledge, et al. 2021. FINQA: A Dataset of Numerical Reasoning over Financial Data. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3697–3711. Association for Computational Linguistics.

Bosheng Ding, Chengwei Qin, Linlin Liu, Yew Ken Chia, Boyang Li, Shafiq Joty, and Lidong Bing. 2023. Is GPT-3 a Good Data Annotator? In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11173–11195, Toronto, Canada. Association for Computational Linguistics.

Saumya Gandhi, Ritu Gala, Vijay Viswanathan, Tongshuang Wu, and Graham Neubig. 2024. Better Synthetic Data by Retrieving and Transforming Existing Datasets. *arXiv preprint arXiv:2404.14361*.

Tao Ge, Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. 2025. Scaling Synthetic Data Creation with 1,000,000,000 Personas. *arXiv preprint arXiv:2406.20094*.

Jianping Gou, Baosheng Yu, Stephen J Maybank, and Dacheng Tao. 2021. Knowledge Distillation: A survey. *International Journal of Computer Vision*, 129(6):1789–1819.

Chetan Harsha, Karmvir Singh Phogat, Sridhar Dasaratha, Sai Akhil Puranam, and Shashishekar Ramakrishna. 2025. Synthetic data generation using large language models for financial question answering. In *Proceedings of the Joint Workshop of the 9th Financial Technology and Natural Language Processing (FinNLP), the 6th Financial Narrative Processing (FNP), and the 1st Workshop on Large Language Models for Finance and Legal (LLMFinLegal)*, pages 76–95, Abu Dhabi, UAE. Association for Computational Linguistics.

Namgyu Ho, Laura Schmid, and Se-Young Yun. 2023. Large Language Models Are Reasoning Teachers. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14852–14882, Toronto, Canada. Association for Computational Linguistics.

Yechan Hwang, Jinsu Lim, Young-Jun Lee, and Ho-Jin Choi. 2023. Augmentation for Context in Financial Numerical Reasoning over Textual and Tabular Data with Large-Scale Language Model. In *NeurIPS 2023 Second Table Representation Learning Workshop*.

Shengzhi Li and Nima Tajbakhsh. 2023. SciGraphQA: A Large-Scale Synthetic Multi-Turn Question-Answering Dataset for Scientific Graphs. *arXiv preprint arXiv:2308.03349*.

Zhuoyan Li, Hangxiao Zhu, Zhuoran Lu, and Ming Yin. 2023. Synthetic Data Generation with Large Language Models for Text Classification: Potential and Limitations. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10443–10461, Singapore. Association for Computational Linguistics.

Xianyang Liu, Yilin Liu, Shuai Wang, Hao Cheng, Andrew Estornell, Yuzhi Zhao, and Jiaheng Wei. 2025. AgenticMath: Enhancing LLM Reasoning via Agentic-based Math Data Generation. *arXiv preprint arXiv:2510.19361*.

Haipeng Luo, Qingfeng Sun, Can Xu, Pu Zhao, Jianguang Lou, Chongyang Tao, Xiubo Geng, Qingwei Lin, Shifeng Chen, and Dongmei Zhang. 2023. WizardMath: Empowering Mathematical Reasoning for Large Language Models via Reinforced Evol-Instruct. *arXiv preprint arXiv:2308.09583*.

Lucie Charlotte Magister, Jonathan Mallinson, Jakub Adamek, Eric Malmi, and Aliaksei Severyn. 2023. Teaching Small Language Models to Reason. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 1773–1781, Toronto, Canada. Association for Computational Linguistics.

Ziyang Miao, Qiyu Sun, Jingyuan Wang, Yuchen Gong, Yaowei Zheng, Shiqi Li, and Richong Zhang. 2025. Easy Dataset: A Unified and Extensible Framework for Synthesizing LLM Fine-Tuning Data from Unstructured Documents. *arXiv preprint arXiv:2507.04009*.

Arindam Mitra, Luciano Del Corro, Guoqing Zheng, Shweti Mahajan, Dany Rouhana, Andres Codas, Yadong Lu, Wei ge Chen, Olga Vrousgos, Corby Rosset, Fillipe Silva, Hamed Khanpour, Yash Lara, and Ahmed Awadallah. 2024. AgentInstruct: Toward Generative Teaching with Agentic Flows. *arXiv preprint arXiv:2407.03502*.

Karmvir Singh Phogat, Sai Akhil Puranam, Sridhar Dasaratha, Chetan Harsha, and Shashishekar Ramakrishna. 2024. Fine-tuning Smaller Language Models for Question Answering over Financial Documents. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, Miami, Florida. Association for Computational Linguistics.

Maximilian Schmidt, Andrea Bartezzaghi, and Ngoc Thang Vu. 2024. Prompting-based Synthetic Data Generation for Few-Shot Question Answering. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 13168–13178, Torino, Italia. ELRA and ICCL.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. Self-Instruct: Aligning Language Models with Self-Generated Instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Zirui Wang, Adams Wei Yu, Orhan Firat, and Yuan Cao. 2021. Towards Zero-Label Language Learning. *arXiv preprint arXiv:2109.09193*.

Ian Wu, Sravan Jayanthi, Vijay Viswanathan, Simon Rosenberg, Sina Khoshfetrat Pakazad, Tongshuang Wu, and Graham Neubig. 2024. Synthetic Multimodal Question Generation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 12960–12993, Miami, Florida, USA. Association for Computational Linguistics.

Hai Ye, Mingbao Lin, Hwee Tou Ng, and Shuicheng Yan. 2025. Multi-Agent Sampling: Scaling Inference Compute for Data Synthesis with Tree Search-Based Agentic Collaboration. *arXiv preprint arXiv:2412.17061*.

Jiacheng Ye, Jiahui Gao, Qintong Li, Hang Xu, Jiangtao Feng, Zhiyong Wu, Tao Yu, and Lingpeng Kong. 2022. ZeroGen: Efficient Zero-shot Learning via Dataset Generation. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 11653–11669, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Longhui Yu, Weisen Jiang, Han Shi, Jincheng YU, Zhengying Liu, Yu Zhang, James Kwok, Zhenguo Li, Adrian Weller, and Weiyang Liu. 2024. MetaMath: Bootstrap Your Own Mathematical Questions for Large Language Models. In *The Twelfth International Conference on Learning Representations*.

Ziqiang Yuan, Kaiyuan Wang, Shoutai Zhu, Ye Yuan, Jingya Zhou, Yanlin Zhu, and Wenqi Wei. 2024. FinLLMs: A Framework for Financial Reasoning Dataset Generation with Large Language Models. *arXiv preprint arXiv:2401.10744*.

Xinyi Zheng, Douglas Burdick, Lucian Popa, Xu Zhong, and Nancy Xin Ru Wang. 2021. Global Table Extractor (GTE): A Framework for Joint Table Identification and Cell Structure Recognition Using Visual Context. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 697–706.

Fengbin Zhu, Wenqiang Lei, Youcheng Huang, Chao Wang, Shuo Zhang, Jiancheng Lv, Fuli Feng, and Tat-Seng Chua. 2021. TAT-QA: A Question Answering Benchmark on a Hybrid of Tabular and Textual Content in Finance. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3277–3287. Association for Computational Linguistics.

A Synthetic Data Generation Prompts

The Content Extraction Agent prompt in Figure 2 guides the agent in extracting text and tables from each PDF page image and in determining the page's structural complexity. Then the Content Selection Agent prompt in Figure 3 is used for evaluating the extracted page content and determining whether a financial reasoning question can be generated. Furthermore, Figure 4 presents the prompt used to guide the agent in formulating financial reasoning questions based on the extracted text and tables, while adhering to predefined constraints and incorporating reflection feedback. The Question Reflection Agent prompt in Figure 5 is used to instruct the agent responsible for evaluating generated questions and providing refinement feedback based on the page content.

B Examples

The following examples illustrate representative samples of both valid and invalid simple and complex pages. Figure 6 shows a valid simple page that includes a single table and accompanying text, making it suitable for generating financial-reasoning questions. Figures 7, 8, 9, and 10 depict invalid simple pages—such as index pages or tables of contents—that the Content Selection Agent filters out because they do not support financial question generation. Figure 11 presents a valid complex page containing multiple tables and relevant numerical information, which makes it appropriate for creating financial-reasoning questions. In contrast, Figures 12, 13, and 14 show invalid complex pages where tables contain more than 20 rows, have merged headers, or lack meaningful financial data necessary for generating insightful questions.

Without the Content Selection Agent, pages like those shown in Figure 15 would be passed directly to the Question Generation Agent. This often results in irrelevant questions, such as "What is the total number of pages dedicated to the Consolidated Financial Statements section?", which are not meaningful in a financial context and can negatively affect SLM performance.

System Prompt

You are an expert in parsing PDF documents. You are tasked with extracting the contents of the page.

User Prompt

Extract the contents of the page from the image of the PDF page provided. Extract the text exactly as it appears in the image.
Extract any table(s) found on the page using the following format:

1. Table headers should be separated by a pipe (|),
2. Table rows should be separated by a newline character (\n).

After extraction, determine the page_type as either "simple" or "complex" based on the following conditions:
A page is classified as complex if it satisfies any of the following:

- Contains more than one table,
- Contains a table with more than 20 rows,
- Contains more than 2 description headers,
- Contains a table with complex nested structures,
- Contains a table with tedious contents, such as catalogs.

Otherwise, classify the page as simple.

If the page is complex, also provide a brief explanation of why it was marked complex. If it is simple, set the reason to null.
The final response should be formatted as a JSON object with the following fields and no others:

{
    "Extracted_data": "<Extracted Text and Table>",
    "Is_Table_Present": "<Yes/No>",
    "page_type": "<simple/complex>",
    "complexity_reason": "<Reason if complex, otherwise null>"
}

Figure 2: Content Extraction Agent Prompt

System Prompt

You are an expert in analyzing parsed document data. You are tasked with determining whether financial reasoning questions can be generated from the extracted page content.

User Prompt

You are provided with the extracted content and metadata from the previous agent, as shown below:
<Extracted_content_from_previous_agent>
Using this extracted content, evaluate whether a multistep financial reasoning question involving arithmetic or logical operations can be framed.
The final response should be formatted as a JSON object with the following fields and no others:

{
    "Can_Financial_Reasoning_Question_Be_Framed": "<True/False>"
}

Figure 3: Content Selection Agent Prompt

System Prompt

You are an expert financial analyst skilled in generating questions that are meaningful for financial analysis or generating insights from the company financial reports. Your task involves analyzing text and table content of a single page from a financial report to generate a relevant question. While performing this task, you must adhere to a set of specified constraints. Use the reflection agent's feedback to improve the quality, clarity, and constraint adherence of the generated question.

User Prompt

Given the text and table content of a page from a financial report, generate a financial reasoning question. Please adhere to the following constraints meticulously when formulating the question.
Text and Table Content: { `text and table content` }
Reflection Feedback to incorporate: { `reflection_feedback` }
Constraints:

1. The question must be generated such that it always leads to a single numerical or boolean answer.
2. The question must be strictly derived from the content present in the image.
3. Answering the question should require the use of values from the table and/or values present in the text around the table.
4. Calculating the answer should involve multi-hop reasoning with the following arithmetic operations:
    - Basic operations: Addition, Subtraction, Multiplication, Division, Exponential, Greater Than
    - Table aggregation operations: Sum, Average, Minimum, Maximum

The final response should be formatted as a JSON object with only question and no other objects should be included:

{
    "Question": "<Generated Question>"
}

Figure 4: Question Generation Agent Prompt

System Prompt

You are an expert financial analyst acting as a reflection agent. Your task is to analyze the generated question in the context of the provided text and table content. You must evaluate the question and provide feedback that helps improve the quality, clarity, and constraint adherence of the generated question.

User Prompt

You are given the following:
Text and Table Content: { `text_and_table_content` }
Generated Question: { `generated_question` }
Review the question carefully in relation to the provided content. Identify any issues related to quality, clarity, relevance, reasoning depth, or adherence to the intended task. Provide feedback that can be directly used to improve the question. Respond ONLY with a JSON object in the following format:
{
"ReflectionFeedback": "<Your feedback here>"
}

Figure 5: Question Reflection Agent Prompt

Chipotle Mexican Grill, Inc.
Notes to Financial Statements (Continued)
(in thousands, except per share data)

earned. Initial fees are recognized upon opening a restaurant, which is when the Company has performed substantially all initial services required by the franchise arrangement.

Cash and Cash Equivalents

The Company considers all highly liquid investment instruments purchased with an initial maturity of three months or less to be cash equivalents.

Accounts Receivable

Accounts receivable consists of tenant improvement receivables, credit card receivables, and miscellaneous receivables. The allowance for doubtful accounts is the Company's best estimate of the amount of probable credit losses in the Company's existing accounts receivable based on a specific review of account balances. Account balances are charged off against the allowance after all means of collection have been exhausted and the potential for recoverability is considered remote.

Inventory

Inventory, consisting principally of food, beverages, and supplies, is valued at the lower of first-in, first-out cost or market. The Company has no minimum purchase commitments with its vendors. The Company purchases certain key ingredients (steak, chicken, pork and tortillas) from a small number of suppliers.

Leasehold Improvements, Property and Equipment

Leasehold improvements, property and equipment are stated at cost. Internal costs clearly associated with the acquisition, development and construction of a restaurant are capitalized. Expenditures for major renewals and improvements are capitalized while expenditures for minor replacements, maintenance and repairs are expensed as incurred. Depreciation is calculated using the straight-line method over the estimated useful lives of the assets. Leasehold improvements are amortized over the shorter of the lease term, which generally includes reasonably assured option periods, or the estimated useful lives of the assets. Upon retirement or disposal of assets, the accounts are relieved of cost and accumulated depreciation and the related gain or loss is reflected in earnings.

The estimated useful lives are:

Leasehold improvements and buildings	3-20 years
Furniture and fixtures	3-10 years
Equipment	3-7 years

Goodwill

Goodwill represents the excess of cost over fair value of net assets of the business acquired. Goodwill resulted from McDonald's purchases of the Company. Goodwill determined to have an indefinite life is not subject to amortization, but instead is tested for impairment at least annually in accordance with the provision of Statement of Financial Accounting Standard ("SFAS") No. 142, *Goodwill and Other Intangible Assets* ("SFAS 142"). In accordance with SFAS 142, the Company is required to make any necessary impairment adjustments. Impairment is measured as the excess of the

Figure 6: Valid simple page

TABLE OF CONTENTS

PART I

Item 1.	Business	1
Item 1A.	Risk Factors	6
Item 2.	Properties	20
Item 3.	Legal Proceedings	21
Item 4.	Submission of Matters to a Vote of Security Holders	21

PART II

Item 5.	Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities	22
Item 6.	Selected Consolidated Financial Data	23
Item 7.	Management's Discussion and Analysis of Financial Condition and Results of Operations	25
Item 7A.	Quantitative and Qualitative Disclosures About Market Risk	38
Item 8.	Financial Statements and Supplementary Data	40
Item 9.	Changes in and Disagreements With Accountants on Accounting and Financial Disclosure	62
Item 9A.	Controls and Procedures	62
Item 9B.	Other Information	62

PART III

Item 10.	Directors and Executive Officers of the Registrant	63
Item 11.	Executive Compensation	63
Item 12.	Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters	63
Item 13.	Certain Relationships and Related Transactions	63
Item 14.	Principal Accounting Fees and Services	63

PART IV

Item 15.	Exhibits Financial Statement Schedules	64
Signatures		65

Figure 7: Invalid simple page– No financial question can be framed

Item 8. Consolidated Financial Statements and Supplementary Data

	Page
Reports of Independent Registered Public Accounting Firm	48
Consolidated statements of income	50
Consolidated statements of comprehensive income	51
Consolidated balance sheets	52
Consolidated statements of cash flows	54
Consolidated statements of shareholders' equity	55
Notes to consolidated financial statements	56
Supplementary data	
Quarterly financial data (unaudited)	109

47

Figure 8: Invalid simple page– No financial question can be framed

INDEX TO EXHIBITS
[Items 15(a)(3) and 15(c)]

Exhibit No.	Description
10.33*	Form of Notice of Grant Award of Restricted Stock Units and Restricted Stock Unit Agreement (Transition Agreement) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.7 to the Registrant's Quarterly Report on Form 10-Q for the quarter ended March 29, 2015 (File No. 1-9183))
10.34*	Form of Notice of Grant Award of Restricted Stock Units and Restricted Stock Unit Agreement (Deferred) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.8 to the Registrant's Quarterly Report on Form 10-Q for the quarter ended March 29, 2015 (File No. 1-9183))
10.35*	Form of Notice of Grant Award of Stock Appreciation Rights and Stock Appreciation Rights Agreement of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.9 to the Registrant's Quarterly Report on Form 10-Q for the quarter ended March 29, 2015 (File No. 1-9183))
10.36*	Form of Executive Severance Plan between the Registrant and each of Messrs. Hund, Jones, Levatich and Olin
10.37*	Form of Transition Agreement between the Registrant and each of Messrs. Levatich and Olin (incorporated herein by reference to Exhibit 10.4 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2009 (File No. 1-9183))
10.38*	Transition Agreement between the Registrant and Mr. Hund dated November 30, 2009 (incorporated herein by reference to Exhibit 10.5 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2009 (File No. 1-9183))
10.39*	Form of Aircraft Time Sharing Agreement between the Registrant and each of Messrs. Levatich, Olin, Jones and Hund and Madame Bischmann (incorporated herein by reference to Exhibit 10.1 to the Registrant's Quarterly Report on Form 10-Q for the quarter ended September 30, 2012 (File No. 1-9183))
10.40*	Form of Non-competition and Non-solicitation Agreement between Harley-Davidson Canada LP, Fred Deeley Imports Ltd. and Harley-Davidson Motor Company, Inc., as amended (incorporated herein by reference to exhibit 10.1 to the Registrant's Quarterly Report on Form 10-Q for the quarter ended June 28, 2015 (File No. 1-9183))
10.41*	Form of Notice of Award of Performance Shares and Performance Shares Agreement (Standard) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.43 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2015 (File No. 1-9183))
10.42*	Form of Notice of Award of Performance Share Units and Performance Share Unit Agreement (Standard International) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.44 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2015 (File No. 1-9183))
10.43*	Form of Notice of Award of Performance Shares and Performance Shares Agreement (Transition Agreement) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan (incorporated herein by reference to Exhibit 10.45 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2015 (File No. 1-9183))
10.44*	Harley-Davidson Retiree Insurance Allowance Plan, as amended and restated effective January 1, 2016 (incorporated herein by reference to Exhibit 10.44 to the Registrant's Annual Report on Form 10-K for the year ended December 31, 2016 (File No. 1-9183))
10.45*	Form of Notice of Award of Performance Shares and Performance Share Agreement (Standard) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan first approved for use in February 2017 (incorporated herein by reference to Exhibit 10.1 to the Registrant's Annual Report on Form 10-Q for the quarter ended March 26, 2017 (File No. 1-9183))
10.46*	Form of Notice of Award of Performance Shares and Performance Share Agreement (Standard International) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan first approved for use in February 2017(incorporated herein by reference to Exhibit 10.2 to the Registrant's Annual Report on Form 10-Q for the quarter ended March 26, 2017 (File No. 1-9183))
10.47*	Form of Notice of Award of Performance Shares and Performance Share Agreement (Transition Agreement) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan first approved for use in February 2017 (incorporated herein by reference to Exhibit 10.3 to the Registrant's Annual Report on Form 10-Q for the quarter ended March 26, 2017 (File No. 1-9183))
10.48*	Form of Notice of Award of Performance Shares and Performance Share Agreement (Special Retention) of Harley-Davidson, Inc. under the Harley-Davidson, Inc. 2014 Incentive Stock Plan first approved for use in February 2017 (incorporated herein by reference to Exhibit 10.4 to the Registrant's Annual Report on Form 10-Q for the quarter ended March 26, 2017 (File No. 1-9183))
21	List of Subsidiaries
23	Consent of Independent Registered Public Accounting Firm
31.1	Chief Executive Officer Certification pursuant to Rule 13a-14(a)
31.2	Chief Financial Officer Certification pursuant to Rule 13a-14(a)

* Represents a management contract or compensatory plan, contract or arrangement in which a director or named executive officer of the Company participated.

117

Figure 9: Invalid simple page– No financial question can be framed

10.98+	First Amendment to the Synchrony Financial Restoration Plan (incorporated by reference to Exhibit 10.118 to 2015 Annual Report on Form 10-K filed by Synchrony Financial on February 25, 2016)
10.99+	Second Amendment to the Synchrony Financial Restoration Plan (incorporated by reference to Exhibit 10.119 to 2015 Annual Report on Form 10-K filed by Synchrony Financial on February 25, 2016)
10.100+	Form of Synchrony Financial Change in Control Severance Plan (incorporated by reference to Exhibit 10.3 to Form 8-K filed by Synchrony Financial on May 27, 2015)
10.101+	Synchrony Financial Amended and Restated 2014 Long-Term Incentive Plan (incorporated by reference to Exhibit 10.1 to Form 10-Q filed by Synchrony Financial on July 28, 2017)
10.102†	Services Agreement, dated as of September 15, 2015, between Retail Finance Servicing, LLC and First Data Resources, LLC (incorporated by reference to Exhibit 10.1 to Form 8-K filed by Synchrony Financial on September 15, 2015)
10.103	Letter, dated as of October 19, 2015, delivered by General Electric Capital Corporation and acknowledged and agreed to by General Electric Company and Synchrony Financial (incorporated by reference to Exhibit 10.116 of Form S-4 Registration Statement filed by Synchrony Financial on October 19, 2015 (No. 333-207479))
10.104+	Amended and Restated form of agreement for awards of Restricted Stock Units and Non-Qualified Stock Options under Synchrony 2014 Long-Term Incentive Plan (incorporated by reference to Exhibit 10.1 to Form 10-Q filed by Synchrony Financial on April 26, 2018)
10.105+	Amended and Restated form of agreement for awards of Performance Share Units under Synchrony 2014 Long-Term Incentive Plan (incorporated by reference to Exhibit 10.2 to Form 10-Q filed by Synchrony Financial on April 26, 2018)
10.106+	Form of agreement for awards of Restricted Stock Units under Synchrony 2014 Long-Term Incentive Plan to directors of Synchrony Financial (incorporated by reference to Exhibit 10.3 to Form 10-Q filed by Synchrony Financial on April 26, 2018)
10.107+	Amended and Restated Executive Severance Plan (incorporated by reference to Exhibit 10.4 to Form 10-Q filed by Synchrony Financial on April 26, 2018)
10.108	Amended and Restated Master Indenture, dated as of May 1, 2018, between Synchrony Card Issuance Trust, as Issuer and The Bank of New York Mellon, as Indenture Trustee (incorporated by reference to Exhibit 4.1 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.109	Form of Class A Terms Document, between Synchrony Card Issuance Trust and The Bank of New York Mellon (incorporated by reference to Exhibit 4.3 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.110	Form of Class B Terms Document, between Synchrony Card Issuance Trust and The Bank of New York Mellon (incorporated by reference to Exhibit 4.4 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.111	Form of Class C Terms Document, between Synchrony Card Issuance Trust and The Bank of New York Mellon (incorporated by reference to Exhibit 4.5 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.112	Form of Class D Terms Document, between Synchrony Card Issuance Trust and The Bank of New York Mellon (incorporated by reference to Exhibit 4.6 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.113	Amended and Restated Trust Agreement, among Synchrony Card Funding, LLC, Citibank, N.A. and Citicorp Trust Delaware, National Association (incorporated by reference to Exhibit 4.7 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.114	Custody and Control Agreement, dated as of November 17, 2017, by and among The Bank of New York Mellon, in its capacity as Custodian and in its capacity as Indenture Trustee, and Synchrony Card Issuance Trust (incorporated by reference to Exhibit 4.8 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.115	Amended and Restated Receivables Sale Agreement, dated as of May 1, 2018, between Synchrony Bank and Synchrony Card Funding, LLC (incorporated by reference to Exhibit 4.9 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))
10.116	Amended and Restated Transfer Agreement, dated as of May 1, 2018, between Synchrony Card Funding, LLC and Synchrony Card Issuance Trust (incorporated by reference to Exhibit 4.10 of Form SF-3 Registration Statement filed by Synchrony Card Issuance Trust and Synchrony Card Funding, LLC on May 4, 2018 (No. 333-224689 and 333-224689-01))

Figure 10: Invalid simple page– No financial question can be framed

Debt Securities

The following discussion provides supplemental information regarding our debt securities portfolio. All of our debt securities are classified as available-for-sale and are held to meet our liquidity objectives or to comply with the Community Reinvestment Act. Debt securities classified as available-for-sale are reported in our Consolidated Statements of Financial Position at fair value.

The following table sets forth the amortized cost and fair value of our portfolio of debt securities at the dates indicated:

	2018		2017		2016	
At December 31 ($ in millions)	Amortized Cost	Estimated Fair Value	Amortized Cost	Estimated Fair Value	Amortized Cost	Estimated Fair Value
Debt:						
U.S. government and federal agency	$ 2,889	$ 2,888	$ 2,419	$ 2,416	$ 3,676	$ 3,676
State and municipal	50	48	44	44	47	46
Residential mortgage-backed	1,180	1,139	1,258	1,231	1,400	1,373
Asset-backed	1,988	1,985	781	780	—	—
U.S. corporate debt	2	2	2	2	—	—
Total	$ 6,109	$ 6,062	$ 4,504	$ 4,473	$ 5,123	$ 5,095

Unrealized gains and losses, net of the related tax effect, on available-for-sale debt securities that are not other-than-temporarily impaired are excluded from earnings and are reported as a separate component of comprehensive income (loss) until realized. At December 31, 2018, 2017 and 2016, our debt securities had gross unrealized gains of $1 million, $1 million and $3 million, respectively, and gross unrealized losses of $48 million, $32 million and $31 million, respectively.

Our debt securities portfolio had the following maturity distribution at December 31, 2018.

($ in millions)	Due in 1 Year or Less	Due After 1 through 5 Years	Due After 5 through 10 Years	Due After 10 years	Total
Debt:					
U.S. government and federal agency	$ 2,888	$ —	$ —	$ —	$ 2,888
State and municipal	—	—	5	43	48
Residential mortgage-backed	1	—	158	980	1,139
Asset-backed	1,613	372	—	—	1,985
U.S. corporate debt	2	—	—	—	2
Total(1)	$ 4,504	$ 372	$ 163	$ 1,023	$ 6,062
Weighted average yield(2)	2.4%	2.7%	3.3%	2.9%	2.5%

(1) Amounts stated represent estimated fair value.
(2) Weighted average yield is calculated based on the amortized cost of each security. In calculating yield, no adjustment has been made with respect to any tax-exempt obligations.

At December 31, 2018, we did not hold investments in any single issuer with an aggregate book value that exceeded 10% of equity, excluding obligations of the U.S. government.

Figure 11: Valid complex page

Exhibit Number	Exhibit Description
10.5	Termination Amendment to Amended and Restated Consulting and Noncompete Agreement of Timothy P. Smucker, dated as of April 25, 2011*
10.6	Termination Amendment to Amended and Restated Consulting and Noncompete Agreement of Richard K. Smucker, dated as of April 25, 2011*
10.7	The J. M. Smucker Company Voluntary Deferred Compensation Plan, Amended and Restated as of December 1, 2012*
10.8	The J. M. Smucker Company 2006 Equity Compensation Plan, effective August 17, 2006*
10.9	The J. M. Smucker Company 2010 Equity and Incentive Compensation Plan*
10.10	Amendment No. 1 to The J. M. Smucker Company 2010 Equity and Incentive Compensation Plan*
10.11	Form of Restricted Stock Agreement*
10.12	Form of Deferred Stock Units Agreement*
10.13	Form of Special One-Time Grant of Restricted Stock Agreement*
10.14	Form of Restricted Stock Agreement*
10.15	Form of Special One-Time Grant of Restricted Stock Agreement*
10.16	Form of Special One-Time Grant of Deferred Stock Units Agreement*
10.17	Form of Restricted Stock Agreement*
10.18	Form of Deferred Stock Units Agreement*
10.19	Form of Performance Units Agreement*
10.20	Form of Restricted Stock Agreement*
10.21	Form of Deferred Stock Units Agreement*
10.22	Form of Nonstatutory Stock Option Agreement*
10.23	Form of Nonstatutory Stock Option Agreement between the Company and the Optionee (three-year vesting)*
10.24	The J. M. Smucker Company Nonemployee Director Deferred Compensation Plan (Amended and Restated Effective January 1, 2007)*
10.25	The J. M. Smucker Company Nonemployee Director Deferred Compensation Plan (Amended and Restated Effective January 1, 2014)*
10.26	The J. M. Smucker Company Defined Contribution Supplemental Executive Retirement Plan, Restated Effective May 1, 2015*
10.27	Amendment No. 1 to The J. M. Smucker Company Defined Contribution Supplemental Executive Retirement Plan, dated as of December 31, 2016*
10.28	The J. M. Smucker Company Restoration Plan, Amended and Restated Effective January 1, 2013*
10.29	Amendment No. 1 to The J. M. Smucker Company Restoration Plan, dated as of May 1, 2015*
10.30	Amendment No. 2 to The J. M. Smucker Company Restoration Plan, dated as of December 31, 2016*
10.31	Form of Amended and Restated Change in Control Severance Agreement between the Company and the Officer party thereto*
10.32	Form of Indemnity Agreement between the Company and the Officer party thereto*
10.33	The J. M. Smucker Company 1998 Equity and Performance Incentive Plan (Amended and Restated Effective June 6, 2005)*
10.34	Tax Matters Agreement between The Procter & Gamble Company, The Folgers Coffee Company, and the Company, dated November 6, 2008
10.35	Intellectual Property Matters Agreement between The Procter & Gamble Company and The Folgers Coffee Company, dated November 6, 2008
10.36	Revolving Credit Agreement, dated as of September 1, 2017, by and among the Company, Smucker Foods of Canada Corp., a federally incorporated Canadian corporation, Bank of America, N.A., as administrative agent, and the several financial institutions from time to time party thereto

Figure 12: Invalid complex page: Contains more than 20 rows but no financial reasoning data

PART IV

Item 15. Exhibits and Financial Statement Schedules.

(a)(1)	Financial Statements:
	See the Index to Financial Statements on page 34 of this Annual Report.
(a)(2)	Financial Statement Schedules:
	Financial statement schedules are omitted because they are not applicable or because the information required is set forth in the Consolidated Financial Statements or notes thereto.
(a)(3)	Exhibits:
	The following exhibits are either attached or incorporated herein by reference to another filing with the U.S. Securities and Exchange Commission.

Exhibit Number	Exhibit Description
2.1	Agreement and Plan of Merger, dated as of February 3, 2015, by and among Blue Acquisition Group, Inc., the Company, SPF Holdings I, Inc., SPF Holdings II, LLC, and, for the limited purposes set forth therein, Blue Holdings I, L.P.
2.2	Stock Purchase Agreement and Plan of Merger, dated as of April 4, 2018, by and among NU Pet Company, PR Merger Sub I, LLC, Ainsworth Pet Nutrition Parent, LLC, CP APN, Inc., CP APN, L.P., and, solely for the limited purpose set forth therein, The J. M. Smucker Company
2.3	First Amendment to Stock Purchase Agreement and Plan of Merger and Side Letter, dated as of May 14, 2018, by and among NU Pet Company, PR Merger Sub I, LLC, Ainsworth Pet Nutrition Parent, LLC, CP APN, Inc., CP APN, L.P., and, solely for the limited purpose set forth therein, The J. M. Smucker Company
3.1	Amended Articles of Incorporation of The J. M. Smucker Company
3.2	Amended Regulations of The J. M. Smucker Company
4.1	Rights Agreement, dated as of May 20, 2009, by and between the Company and Computershare Trust Company, N.A., as rights agent
4.2	Amendment No. 1, dated as of February 3, 2015, to the Rights Agreement, dated as of May 20, 2009, between the Company and Computershare Trust Company, N.A., as rights agent
4.3	Amendment No. 2, dated as of October 24, 2016, to the Rights Agreement, dated as of May 20, 2009, by and between the Company and Computershare Trust Company, N.A., as rights agent
4.4	Amendment No. 3, dated as of June 25, 2018, to the Rights Agreement, dated as of May 20, 2009, by and between the Company and Computershare Trust Company, N.A., as rights agent, and subsequently amended as of February 3, 2015, and October 24, 2016
4.5	Indenture, dated as of October 18, 2011, between the Company and U.S. Bank National Association
4.6	First Supplemental Indenture, dated as of October 18, 2011, among the Company, the guarantors party thereto, and U.S. Bank National Association
4.7	Third Amended and Restated Intercreditor Agreement, dated June 11, 2010, among the administrative agents and other parties identified therein
4.8	Indenture, dated as of March 20, 2015, between the Company and U.S. Bank National Association, as trustee
4.9	First Supplemental Indenture, dated as of March 20, 2015, by and among the Company, the guarantors party thereto and U.S. Bank National Association, as trustee
4.10	Second Supplemental Indenture, dated as of December 7, 2017, between the Company and U.S. Bank National Association, as trustee
10.1	Nonemployee Director Stock Plan dated January 1, 1997*
10.2	The J. M. Smucker Company Top Management Supplemental Retirement Benefit Plan, restated as of January 1, 2018*
10.3	Amended and Restated Consulting and Noncompete Agreement of Timothy P. Smucker, dated as of December 31, 2010*
10.4	Amended and Restated Consulting and Noncompete Agreement of Richard K. Smucker, dated as of December 31, 2010*

Figure 13: Invalid complex page: Contains more than 1 table but no financial reasoning data

Description	Exhibit Number
Amendment No. 1 dated January 1, 2003 to Supplemental Retirement Agreement between Registrant and Jonathan M. Tisch, incorporated herein by reference to Exhibit 10.37 to Registrant's Report on Form 10-K for the year ended December 31, 2002	10.18^{+}
Amendment No. 2 dated January 1, 2004 to Supplemental Retirement Agreement between Registrant and Jonathan M. Tisch, incorporated herein by reference to Exhibit 10.41 to Registrant's Report on Form 10-K for the year ended December 31, 2003	10.19^{+}
Form of Stock Option Certificate for grants to executive officers and other employees and to non-employee directors pursuant to the Loews Corporation Amended and Restated Stock Option Plan, incorporated herein by reference to Exhibit 10.27 to Registrant's Report on Form 10-K for the year ended December 31, 2009	10.20^{+}
Form of Award Certificate for grants of stock appreciation rights to executive officers and other employees pursuant to the Loews Corporation Amended and Restated Stock Option Plan, incorporated herein by reference to Exhibit 10.28 to Registrant's Report on Form 10-K for the year ended December 31, 2009	10.21^{+}
Lease agreement dated November 20, 2001 between 61st & Park Ave. Corp. and Preston R. Tisch and Joan Tisch, incorporated herein by reference to Exhibit 10.1 to Registrant's Report on Form 10-Q filed August 4, 2009	10.22
(21) Subsidiaries of the Registrant	
List of subsidiaries of the Registrant	21.01*
(23) Consent of Experts and Counsel	
Consent of Deloitte & Touche LLP	23.01*
(31) Rule 13a-14(a)/15d-14(a) Certifications	
Certification by the Chief Executive Officer of the Company pursuant to Rule 13a-14(a) and Rule 15d-14(a)	31.01*
Certification by the Chief Financial Officer of the Company pursuant to Rule 13a-14(a) and Rule 15d-14(a)	31.02*
(32) Section 1350 Certifications	
Certification by the Chief Executive Officer of the Company pursuant to 18 U.S.C. Section 1350 (as adopted by Section 906 of the Sarbanes-Oxley Act of 2002)	32.01*
Certification by the Chief Financial Officer of the Company pursuant to 18 U.S.C. Section 1350 (as adopted by Section 906 of the Sarbanes-Oxley Act of 2002)	32.02*

Figure 14: Invalid complex page– No financial question can be framed

Item 8. Financial Statements and Supplementary Data.

Item 9. Changes in and Disagreements with Accountants on Accounting and Financial Disclosure.

None.

Item 9(a). Controls and Procedures.

In accordance with the Securities Exchange Act of 1934 Rules 13a-15 and 15d-15, we carried out an evaluation, under the supervision and with the participation of management, including our Chief Executive Officer and Chief Financial Officer, of the effectiveness of our disclosure controls and procedures as of the end of the period covered by this report. Based on that evaluation, our Chief Executive Officer and Chief Financial Officer concluded that our disclosure controls and procedures were effective as of December 31, 2015 to provide reasonable assurance that information required to be disclosed in our reports filed or submitted under the Exchange Act is recorded, processed, summarized, and reported within the time periods specified in the Securities and Exchange Commission's rules and forms. Our disclosure controls and procedures include controls and procedures designed to ensure that information required to be disclosed in reports filed or submitted under the Exchange Act is accumulated and communicated to our management, including our Chief Executive Officer and Chief Financial Officer, as appropriate, to allow timely decisions regarding required disclosure.

There has been no change in our internal control over financial reporting that occurred during the three months ended December 31, 2015 that has materially affected, or is reasonably likely to materially affect, our internal control over financial reporting.

See page 42 for Management's Report on Internal Control Over Financial Reporting and page 44 for Report of Independent Registered Public Accounting Firm on its assessment of our internal control over financial reporting.

Item 9(b). Other Information.

None.

Figure 15: Example of a page that would be passed to the Question Generation Agent leading to low quality question generation: Content Selection Agent filters such pages