Exploratory Data Analysis and the Rise of Large Language Models - Gaming Industry Insights

Denitsa Zhecheva [1]

[1] *Konstantin Preslavsky University of Shumen, Universitetska str. 115, Shumen, Bulgaria*

Abstract **– The applications of modern large language models are diverse and new at the same time. It is forecasted that the scientific society and businesses may experience a long period of time exploring all opportunities and challenges for using them which will allow analysis of the impact on how advanced generative artificial intelligence is changing work occupation activities and performance efficiency. Undoubtedly, today's question that every business must answer is not if but how to implement large language models, due to their ability to transform numerous business processes. This study aims to give a better understanding on how large language models are contributing to the process of exploratory data analysis as they are not here to replace the traditional methods but to add generative artificial intelligence capabilities to the well-established ones. The results of this paper reveal high level of accuracy of the paired output between operation prompts in OpenAI's large language model and human-mediated entry. However, such output comparison highlighs the need for more informative and specific input prompts to ascertain this accuracy. Further caveats that need to be placed in consideration refer to possible system downtimes, as well as the expenses incurred with every prompt execution. Nevertheless, the comparative speed of operation of large language models remains their most substantial competitive advantage.**

Overall, the findings in this paper contribute to understanding that large language models streamline with ease the desired extraction of insightful information which may further be used for better decision-making, good data management, and design of winning growth strategy as is the case of the gaming industry.

Keywords **– Large language models, exploratory data analysis, gaming industry, generative artificial intelligence, natural language processing.**

DOI: 10.18421/TEM131-59
https://doi.org/10.18421/TEM13-59

Corresponding author: Denitsa Zhecheva,
Konstantin Preslavsky University of Shumen, Universitetska str. 115, Shumen, Bulgaria
Email: mrszhecheva.denitsa@gmail.com

Received: 15 September 2023.
Revised: 12 December 2023.
Accepted: 19 December 2023.
Published: 27 February 2024.

1. Introduction

The introdcution conveys inforamtion about the significant progress of gaming industry backed by recent statistical reports. This part helps to clarify why the author has chosen data from this particular industry to power it's research. The section then continues with the evolution of large language models by paying attention to the most noticeable ones for the last five years and their connection to exploratory data analysis. The aim and objectives of the article are thoroughly reviewed at the end.

1.1. Gaming Industry Overview

Gaming industry has been subject to significant evolution during the last decade. This demonstration of progress can be partially measured by revenues. Gaming market revenue worldwide doubled - from 91.4 billion U.S. dollars in 2015 to 184.4 billion U.S. dollars in 2022 [3]. Mobile gaming, for example, has been a major driver of this growth, thanks to the increasing user penetration rates in developing countries reaching 74 billion U.S. dollars alone in 2020 [4]. Also, the palette of gaming products is enlarging more and more progressively. Starting with virtual reality (VR) and augmented reality (AR) technologies, going through Indie games (short for independent video game), to Free-to-play (F2P) games which have become a dominant monetarization model with 22.7 billion U.S. dollars revenue worldwide in 2020 [4].

Moreover, gaming industry does not only bet on diverse and innovative products, but it is also investing in different ways to access playing. For example, cloud gaming services which are allowing to stream games directly to players' devices, also, cross-platform play, and allowing users to play together on different types of devices. All this diversity has facilitated the distribution of games to a global audience with the main idea of sustaining player engagement.

All aforementioned facts present sound reasons to conclude that playing games nowadays is almost for anyone. Gaming is an activity encompassing various profiles of people – those favoring single player setting and stretching to the other end of the spectrum – users who prefer multiplayer style (inclusive of all socializing functionalities). Gaming industry pays attention to it all and this is one of the main reasons competitive gaming have been advancing and attracting immense audience through gaming tournaments and sponsorship. Additionally, modern day games allow users to both play and generate extra source of income through live streaming platforms and game tutorials. These impressive breakthroughs of gaming industry are mainly connected to the digital era and the fact that the latter is reinventing almost everything in our lives. Board games, for example, have been around for literally thousands of years [5] and they are now transferred online. Board games apps mirror the actions of a physical board games and add digital features which are proven to be attractive for players.

Gaming thrived in technology and hence in usage leading to huge volumes of data that needs to be collected and processed to provide pivot for further growth. Because the insights generated by the accumulated business information are the compass that shows companies the right way to which standards to raise to and what kind of improvements to be made in details. Businesses become wise and proactive about this matter and, as a result, all specific aspects of a play are subject to collection, observation, and analysis. Some of the critical indicators of investigation include but are not limited to: 1) previously ownership of the game; 2) player status ("want-to-play", 'want-to-buy", etc.); 3) game rating (median, average, Bayesian); 4) availability, and sentiment of comments; 5) games genres; 6) minimum number of players requirement; 7) age limit; 8) availability of expansions; 9) preferred time slot for interaction with the platform ("night owl" or daytime player). Overall, companies invest considerable resources into establishing, understanding of consumer behavior and yield important baselines for garnering user data which is later manipulated through exploratory data analysis (EDA).

1.2. LLMs Evolution

The EDA was first mentioned in 1970s by J. Tukey [6] and it is a combination of methods for data exploration which at the end may exhibit important insights from data. Over the last decade the world of data analytics, and especially in the field of EDA, has been experiencing steady progress. One main pillar of knowledge was supporting the course of examination of business data at that period – the knowledge of programming language that has a rich environment of libraries for data processing, data visualization, machine learning modeling, deep learning modeling, etc. All aforementioned well-established libraries along with the expertise and know-how of a strong programming community are the core facilitator supporting companies' process of EDA. Because the more data a company collects, the more exploration of it has to be made and more people have to be involved in it. In 2017 Assoc. Prof. K. R. Srinath concluded that with the boom of big data an insistent demand for Python developers has appeared for taking positions as data scientists [7]. Also, scientific literature targeting Python for data science has witnessed immense demand during and all this comes for a reason. Undoubtedly, Python is one of many programming languages that helps to systematize, visualize, and derive valuable insights from business data, and it is so widely used because it has a huge ecosystem of libraries used for data exploration: Pandas (for data manipulation and analysis); Scikit-learn (for machine learning algorithms); SciPy (for scientific and technical computing); NumPy (for supporting large, multi-dimensional arrays and matrices); Matplotlib (for plotting); Seaborn (for data visualization); Plotly (for graphing) and many more.

In terms of capacity and speed of operation another tendency has been growing gradually together with the already established standards for EDA – natural language processing (NLP). NLP can be explained as a subset of artificial intelligence and linguistic, that is dedicated to helping computers understand words or statements in the way humans do [8]. In terms of data analysis and EDA, the first step was made with the release of "word2vec" algorithm in 2013, supported by Google. The name stands for "word to vector" and it is made up of two learning algorithms: Continuous Bag of Words (CBW) and Skim-gram [9]. They transform the input text into vector and this way a semantic relation can be found. Then in 2017, the Transformer architecture was introduced to the world, based on Long Short-Term Memory (LSTM) deep learning method. It is a novel neural network architecture based on a self-attention mechanism that is well suited for language understanding (presented in "Attention Is All You Need" by Google).

The latter also shows substantial results in academic translations and improvement of indicators like training speed when fitting modern machine learning models [10]. More precisely, the Transformer architecture has its roots back in 1992 when Jürgen Schmidhuber published his alternative to Recurrent Neural Networks (RNNs) – a transformer with linearized self-attention [11]. The modern transformers model afterwards laid the foundations of large language models (LLMs). LLM is a language model using deep leaning algorithms and it is trained on immense amount of text data.

This way the model is able to generate human-like text and have a great performance in tasks like questions answering [12]. The particular order in which most noticeable LLMs were released from 2018 to the first quarter of 2023, with structured information about each one is presented in Figure 1 (author's own elaboration through [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]).

As it is presented on the "Modern LLMs Timeline Diagram", technology-restructuring models have been released for a very short period of time. Two related to EDA products, powered by part of the presented LLMs are highlighted in this paper due to their innovativeness and effectivity in the process of data exploration: OpenAI's API and ChatGPT.

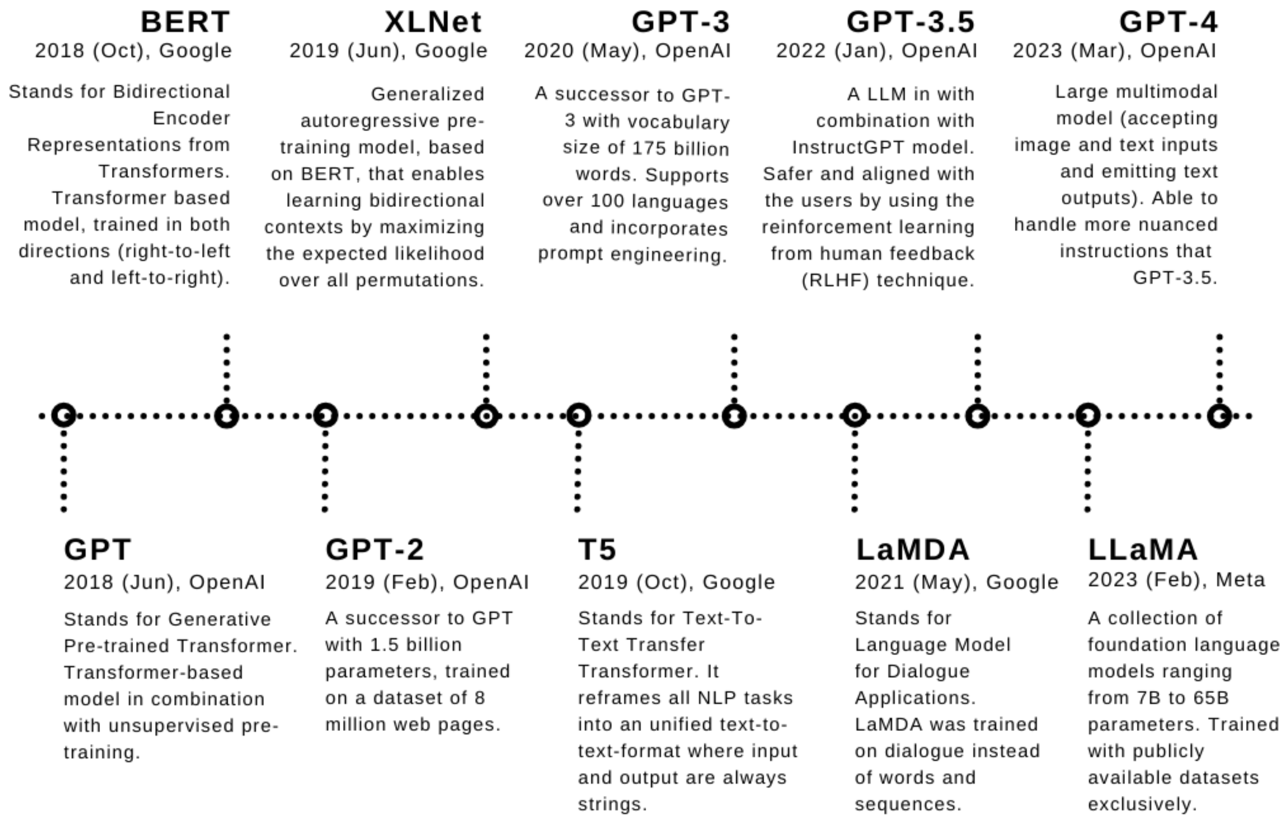

Figure 1. Modern LLMs timeline diagram

Firstly, with the release of the OpenAI's API services in 2020, an option to connect LLM to standard data manipulation libraries for EDA enables was included. Only nine months later, OpenAI reported that over 300 apps across different industries started to use GPT-3 model through the API services: Fable Studio ("virtual beings"), Algolia (neural search and answers products), Viable (data analysis services), Genei pro (ai-generated research tool) and many more.

Also, users started to contribute to the platform and results reached unprecedented levels - 4.5 billion words have been generated on average per day, and the traffic continue to scale [23]. This is where the power of such products comes from – involving humans directly in the process of language fine-tuning.

Next in order, in 2022 it was stated that OpenAI's first commercial product (ChatGPT), powered by the GPT-3.5 model is at everyone's disposal. An ability to talk to the data opens up through the releasing of this large language model-based chatbot.

This new capability to transform the data in a way that Generative Artificial Intelligence (GenAI) takes part in the dialogue of getting insights from the data makes the process of EDA a time-saving and much easier. Slightly after its release, ChatGPT enters and starts contributing to many industries in diverse ways. From an important role in public health through supporting individuals and communities in making informed decision about their well-being [24] and advancing the understanding of climate change, and improving the accuracy of climate projections [25], to improving the independence and autonomy of autodidact learners, [26] and writing law school exams without human assistance [27]. Same process of exploration and analysis of the effects of activating ChatGPT's power appears in gaming industry as it has a long history of using AI tools for game design and creation. Only recently "Game jams, Hackathons, and Game creation events" conference proceedings were published and a valuable conclusion about generating game content was made – the highly accessible ChatGPT offers game jams' participants the opportunity to generate partially or fully an original game by writing only a text prompt [28]. Also, ChatGPT's grading and feedback generation skills have already been evaluated by a team from two USA universities through a comparison with same operations performed by human experts in the context of a digital learning game [29]. Other study describes ChatGPT as preferable colleague in many game development tasks such as programming and 2D art [30]. The last case study mentioned above also found out some weak points like: poor understanding of emotions and empathy, limited creativity and innovativeness which are partially the reasons why ChatGPT is not yet ready to replace full-fledged developer [30]. Thus, another scientific analysis revealed that regardless of its good overall understanding of the poker game, ChatGPT cannot embody ultimate player characteristics due to lack of proper awareness of main concepts like valuation of starting hands, playing positions, and others [31]. Essentially, LLMs have to evolve more to a state where they can comprehend contextual features similar to how humans do.

1.3. Aim and Objectives

The aforementioned limitation of LLMs is considered to be a downside especially in terms of EDA due to its direct relation to the context. As a result, insufficient assistance of the model is underlined later in this paper. Regardless of that, in scientific literature it is believed that LLMs are a useful tool in terms of EDA and this productive interplay may save time and effort [2].

However, most of present research scarcely focuses on the contribution of LLMs to EDA. Also, a lack of measurement indicators to confirm the proposed assumptions about the impact of LLMs on EDA is noted. This article attempts to address the aforementioned gap by conducting scientifically-grounded analysis over time-tested Python libraries for EDA (Pandas, Numpy, Matplotlib, etc.), together with one new Python library (PandasAI) that interacts with recently attracted significant attention – GPT-3.5 model of OpenAI. In order to achieve the aim an examination of the online boardgames market is completed in order to: 1) measure and evaluate the performance of LLMs over diverse data manipulation tasks, forming the process of EDA; 2) distinguish LLMs' positive and negative aspects, domains for improvement. The results of this paper aim to provide better understanding over the implications of LLMs' implementation into the process of EDA through inspecting accuracy, execution time, downtime frequency, etc. The outcome of this research may be beneficial to academia in receiving one more viewpoint over modern LLMs' performance on specific tasks and fulfilling the lack of rich insight that an analysis about supercharging standard EDA libraries' capabilities could provide. It may also contribute to businesses by easening the process of decision-making about what LLM-based products to invest in.

2. Research Methodology

The first part of this section presents an overview of the research methodology through showing attention to the datasets used in the research, evaluation parameters, and applied methods. The following chapter explains in detail the exploration process phase by phase. Following that, all the documented observations of the author on each step are noted.

2.1. Methodological Underpinnings

The process of EDA can be explained as a series of actions to see what information a data can communicate. It is not tied to strictly defined data modelling or formal methodological order for processing, because EDA relies heavily on the data scientist's subjective frame of reference. However, there are several principles which are good to be followed in order to extract the most insights from data: 1). similarity/dissimilarity demonstration; 2). causality demonstration; 3). involvement of multiple variables demonstration.

An EDA that incorporates all the principles mentioned above is being executed by the author and four well distinguished phases are noticed: 1) reading and inspection; 2) plotting distributions, observing, and forming hypotheses; 3) visualizing and confirming/ rejecting of the formed hypotheses; 4). clusters searching, examination and visualization.

Data stemming from a gaming community is used for this study in order to offset some deficiencies of using publicly announced insights by the gaming industry (industry reports, specialized events' agenda, latest news, and gaming-related publications). This is believed to furnish the research with both prime and up-to-minute source of information by delving into online gaming community's preferences and play patterns. That is why, the data inspected by the author, has been collected for the period of January 2022 until February 2023 from an online forum for online board games with a database that includes almost all aspects of players' behavior. The gaming community in question is multinational and it has worldwide representation.

Two datasets form the data inspected in this study and together they represent information about more than 4800 online board games, as well as over 150,000 ratings from over 1,500 users. The first dataset includes detailed information about boardgames such as: title, year of publication, minimum/ maximum players per game, main publisher, category, designer, available expansions, etc. The second one contains more information about their ratings: comments, average rating, user's number of plays, rating's last modification date, whether the user have the game in their wish list, whether the user has previously owned the game, etc. Several Python libraries for data analysis are used for data inspection (Pandas, Matplotlib, Plotly, Numpy, Seaborn), together with PandasAI (0.8.2 version) – Python library that works with GenAI models and, most importantly for the purposes of this research, gives the opportunity to implement an LLM (PandasAI supports only the OpenAI's LLM model and the API services are used in this study). This new tool makes Pandas conversational through sending prompts (user input requests) about the data in natural language. The OpenAI's GPT model is decomposing the text requests into a sequence of words and then predicts a human-like output.

The main sequence of methods applied to this research can be explained as a comparison between traditional libraries for EDA and the AI-powered one through duplicating each operation with different code to invoke the corresponding library. The idea standing behind this approach is, on one hand, to compare the LLM's performance over the traditional methods for EDA, on the other hand, to confirm the truthiness of the generated by the LLM information. In the course of comparing, one quantitative aspect of the performance evaluation (execution time) was measured by implementing a tool that estimates the execution time of each task. The other aspects of the evaluation were qualitative and they were documented as a conclusion after each two matching operations in the working notebook.

2.2. The Process of Exploration

The process of EDA starts with reading the data and then inspecting it (Phase 1). In addition, displaying patterns of missing values is performed and creation of additional columns to improve the upcoming detailed inspection of the data was conducted. This is the place from which the interaction between the LLM and the data starts - the GPT model was asked to inspect the data types of all columns and to suggest improvements (if needed). The LLM model performed well by reading the features' data types correctly and switching part of them. Two types of improvement goals cause this rewriting: 1). memory optimization (from int64 to int32 data types); 2). easier processing and contextualization (from int to bool data types). This operation was executed for 14.7 seconds. At the same time, it took more time to observe and optimize features' data types with traditional methods. The execution time spent for AI-generated task completion was acceptable in this case. However, a lack of ability to understand a related request appeared shortly after when the LLM was asked to return the data types to their original formats. This task was not fulfilled, despite testing with different formulations of the text prompt. Also, a big delay in response (output) was noted (up to 19 seconds per prompt). Additionally, the LLM gave no explanation about this inability (by throwing an error for example). The last-mentioned behavior might become a problematic aspect, when similar tasks are automatically implemented in code, because it will be more difficult to be detected.

The next step of exploring the data (Phase 2), is to plot the entire distribution of all variables. Generating descriptive statistics is one way of achieving this and it was successfully performed by the LLM. Although, it did not manage to fulfill a related task to find some important trends in the presented statistics. It becomes clear that this task is again reserved for data specialists and their ability to spot patterns. An ability which GenAI was initially programmed to unfold in time, but is still not applicable in all cases.

After some dependencies are detected and considered worth examining, a phase for data visualization follows (Phase 3).

The act of visualization of some data chunks occurs on almost every stage of EDA as humans understand information better this way and it is easier to communicate the data with people being involved with the research. On this particular stage, plotting different views of the data is leading to: confirming observations or proving to be wrong; detecting outlines and unusually formed groups; determining clusters; recognizing trends; etc. Histograms, boxplots, pie charts, scatterplots, and a correlation matrix (Figure 2) were plotted by the LLM, in comparison with creating the same visualizations via Seaborn, Matplotlib, and Plotly libraries. The GPT model mapped out the requested data projections in precise and in speed, with exceptionally good timing of operations execution. It must be noted that the LLM reveals its capacity in this task by plotting data into its own consideration. The understanding of the data in such tasks is undoubtedly one of the features that would mostly be valued when choosing a GenAI-based product for EDA. Because it imitates successfully idea generation and problem-solving processes by communicating another perspective of data visualization.

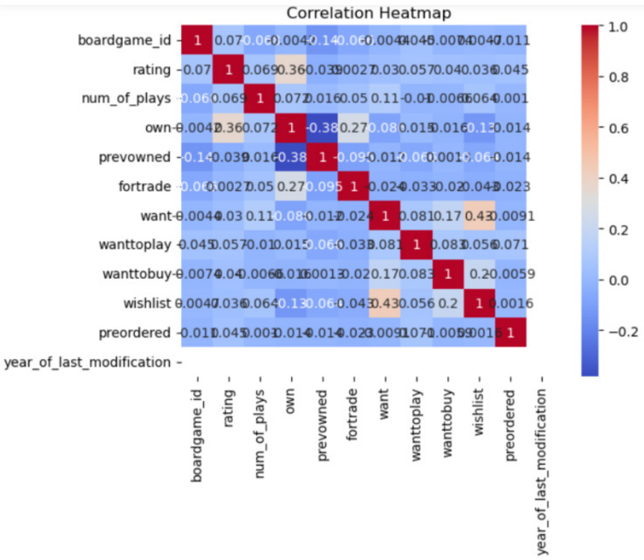

Figure 2. Correlation heatmap generated through prompt to the large language model

Next, the EDA is heading to applying grouping, averaging, summing, and other similar operations (Phase 4) over groups of data to get the idea of possible clusters. The GPT model is considered advanced and speedy on these types of operations and the study demonstrates that it took only several microseconds for execution of the above-mentioned operations. For example, displaying requested series of data which later would be a subject to filtering (grouping operation). Filtering is another advancement in terms of performance of the GPT model.

Questions like "At what time of the day most of the reviews were received?" and commands such as "Filter the DataFrame by comments that are not None" only took up to 1 second to be answered and obtained with accuracy.

Additionally, a specific behavior of the LLM is observed in this phase of the EDA. It is not related to the mentioned-above operations for grouping, averaging, summing, or filtering and requires complex understanding of the data. The question: "In which 4-hours slot of time the majority of ratings were given?" a relevant output was produced ('The majority of ratings were given between 04:00 and 7:59'). In such questions, a semantic involvement is marked. The question that arises is whether the answer is a result of good semantic understanding of the data by the LLM? The author believers that the reason a proper output was received in this task is because of the more informative prompt which, in this case, directs the LLM to observe a specific data type – DateTime. This structure of the prompt facilitated efficient task completion. This paper pays attention to another similar task. A prompt associated with players' nature and habits was sent to the LLM: "Are users night owls or daytime players?". In 8 out of 9 attempts to formulate the right interpretation of the request and to result in receiving a passable output, the LLM did not understand properly the question. Only one of the prompts was understood partially by the LLM and it created series of data, showing the number of ratings given at all hours of the day (grouping operation). Based on this output, data scientists could manage to recognize some patterns and make a conclusion. The author went through this process – online boardgame players are both "night owls" and daytime players, with a small prevail of "night owls" Thus, a further analysis must be made to reveal the specific characteristics of both groups of players. A similar output was expected by the LLM and this is one of the domains that an improvement is expected. Also, if a wider observation to the interpretation of the question that the LLM understood is made ("Are players night owls or daytime players due to the time of day they left their reviews?"), another conclusion arises. This prompt is more specific than the other eight formulation and it directs the LLM "to look up" into a specific data. This way, a slight semantic hint is given to the LLM and it again resulted in acceptable output.

3. Results and Discussion

GenAI is playing a central role in data science. LLMs are its direct representors and one of the related statements that circulates through scientific literature is that an inaccuracy of the extracted knowledge is detected.

In the case of this research, such declaration cannot be made. Through the process of data exploration, multiple duplicated operations were executed by the author with the purpose to verify the LLM's output accuracy. As a result, a *100%* ***accuracy*** was spotted – all outputs, received by tasks fulfilled by human action matched the LLM-generated ones. The reason standing behind this success rate might be relevant to the prompt structured, which includes the inspected DataFrame as part of every request.

Another conclusion arises in the process of research, which again relates to previous findings in scientific literature - there is still room for improvement due to ***semantic understanding*** of the data by the LLMs. This research finds that a semantic understanding is unveiled when the user input request is more informative and specific and this way an accurate output is received. It can be stated that the more information the LLM receives through user's prompt, the better will be the output. This finding could be flagged as best practice when composing text prompts.

Prompt structure happen to be one of the components to consider when discussing ***execution time*** of tasks powered by the OpenAI' GPT model. Also, the type of the requested operation matters. The study finds that LLMs are speedy in the following operations, but not limited to: reading, overviewing, displaying missing values, filtering, grouping, summing, and averaging. Also, the LLM needs only up to 1.06 seconds for plotting and visualizing all the diverse chunks of data – correlations and distributions of different variables. However, a delay in receiving some outputs is reported during the research. It can be concluded that the more time the output takes, the bigger is the possibility for operation unfulfillment. This finding refers again to the need of improvement of the semantic understanding by the LLM, because the mentioned delays were almost 100% requiting it.

As every emerging technology, there is also room for improvement due to ***downtime and partial outages***. During the process of research, the author witnessed two partial outages and this resulted in work flow interruption. When such event occurs, the official website of OpenAI is of service as there is 24-hour reposting system about the condition of the services they offer. Temporary unavailability may be acceptable for personal research purposes, but when it comes to businesses such events shall be considered with discretion.

Another consideration that has to be made before implementing the LLM into one's business is the ***expense*** for using the chosen technology.

The pricing for using the OpenAI's API services, powered by the GPT-3.5-Turbo model, as of May 2023 is approximately 0.02 USD for 1000 tokens. It is important to notice that with every input request by the user the whole DataFrame, along with the request is being processed by the LLM. Companies are processing huge amount of data on a daily basis and a request optimization shall be taken into account.

One more process to look at due to the combining of LLMs' and existing EDA tools is ***automation***. It is always a best business practice to have some processes automated as this saves time and effort. Regarding this optimization, many data analysts report that repetitive tasks are impeding the process of data exploration (they have to execute similar tasks in each EDA) and that most of the traditional tools do not incorporate such knowledge [1]. The implementation of an LLM-based tool for EDA would solve this exploration challenge and let experts focus on other problematic aspects. One consideration that has to be taken preemptively is observed in this study – a lack of information by the LLM about an unfulfilled tasks is observed in some cases. This behavior should be considered when automation planning takes place.

4. Conclusion

This paper has marked important insights on how large language models can contribute to the process of exploratory data analysis. After 228 duplicate operations performed by the author, their beneficial quality was revealed through parameters like processing speed, maximum output accuracy, and optional automation. However, parameters such as downtime, semantic understanding, and potential expenditures fail to meet the standards for excellence. Consequently, these shortcomings might continue to be actively tested in future as businesses are seeking to unlock the full potential of large language models. Building upon the present study, there are even more future directions for research such as: increasing the number and variety of duplicated operations, including more large language models in the comparison, etc. This study could serve as an example of best practices for data analysts, data scientists, and artificial intelligence engineers contributing with their work to various business industries.

To summarize, though the combination of company-generated data and advanced technologies like AI-powered chatbots and LLM-based API services, a more powerful business expansion can be witnessed, as is the case with gaming industry.

Certainly, there are considerations to be taken, but an implementation of GenAI-based tools for EDA is an imperative in order to scale and be major player in industry.

Acknowledgements

This paper is financed by project № РД-08-125/ 22.02.2023 "Parallel algorithms for data protection in graphical objects and three-dimensional modelling".

References:

[1]. Wongsuphasawat, K., Liu, Y., & Heer, J. (2019). Goals, process, and challenges of exploratory data analysis: An interview study. *arXiv preprint arXiv:1911.00568*.

[2]. Amer-Yahia, S., Bonifati, A., Chen, L., Li, G., Shim, K., Xu, J., & Yang, X. (2023). From Large Language Models to Databases and Back: A discussion on research and education. *arXiv preprint arXiv:2306.01388*.

[3]. Clement, J. (2023). *"Games market revenue worldwide from 2015 to 2022, by region".* Retrieved from: https://www.statista.com/statistics/539572/games-market-revenue-by-region/ [accessed: 10 August 2023].

[4]. Clement, J. (2022). *"Digital games industry revenue worldwide in 2020, by game category".* Retrieved from: https://www.statista.com/statistics/379682/digital-games-revenue-distribution-category-usa/ [accessed: 13 August 2023].

[5]. Zwang, D. (2017). *Senet and Twenty Squares: Two Board Games Played by Ancient Egyptians.* metmuseum.org. Retrieved from: https://www.metmuseum.org/blogs/metkids/2017/ancient-egypt-board-games [accessed: 05 September 2023].

[6]. Tukey, J. W. (1977). *Exploratory Data Analysisx.* Addison-Wesley.

[7]. Srinath, K. R. (2017). Python–the fastest growing programming language. *International Research Journal of Engineering and Technology*, *4*(12), 354-357.

[8]. Chopra, A., Prashar, A., & Sain, C. (2013). Natural language processing. *International journal of technology enhancements and emerging engineering research*, *1*(4), 131-134.

[9]. Ma, L., & Zhang, Y. (2015, October). Using Word2Vec to process big text data. In *2015 IEEE International Conference on Big Data (Big Data),* 2895-2897. IEEE.

[10]. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

[11]. Schmidhuber, J. (1992). Learning to control fast-weight memories: An alternative to dynamic recurrent networks. *Neural Computation*, *4*(1), 131-139.

[12]. Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

[13]. Devlin, J., Chang, M.-W., Lee, K. &Toutanova, K. T. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics*, 4171–4186.

[14]. Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding with unsupervised learning.* Technical Report, OpenAI. Retrived from: https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf [accessed: 09 September 2023].

[15]. Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, *32*.

[16]. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

[17]. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., ... & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, *21*(1), 5485-5551.

[18]. Collins, E., & Ghahramani, Z. (2021). LaMDA: our breakthrough conversation technology. *Google AI Blog*.

[19]. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, *33*, 1877-1901.

[20]. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., ... & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*, 27730-27744.

[21]. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

[22]. OpenAI. (2023). GPT-4 Technical Report. *arXiv preprint arXiv:2303.08774*

[23]. OpenAI & A. Pilipiszyn. (2021). GPT-3 powers the next generation of apps. *OpenAI Blog.* Retrived from: https://openai.com/blog/gpt-3-apps [accessed: 11 September 2023].

[24]. Biswas, S. S. (2023). Role of chat gpt in public health. *Annals of biomedical engineering*, *51*(5), 868-869.

[25]. Biswas, S. S. (2023). Potential use of chat gpt in global warming. *Annals of biomedical engineering*, *51*(6), 1126-1127.

[26]. Biswas, S. (2023). Role of Chat GPT in Education. *Journal of ENT Surgery Research, 1*, 1–3.

[27]. Choi, J., Hickman, K., Monahan, A., Schwarcz, D. (2023). ChatGPT Goes to Law School. *Journal of Legal Education.*

[28]. Grow, A. M., & Khosmood, F. (2023, August). ChatGPT GameJam: Unleashing the power of Large Language Models for Game Jams. In *Proceedings of the 7th International Conference on Game Jams, Hackathons and Game Creation Events*, 51-54.

[29]. Nguyen, H. A., Stec, H., Hou, X., Di, S., & McLaren, B. M. (2023, August). Evaluating chatgpt's decimal skills and feedback generation in a digital learning game. In *European Conference on Technology Enhanced Learning* 278-293. Cham: Springer Nature Switzerland.

[30]. Kings, M. & Täcklind. S. (2023). *Leveling Up the Playing Field: Exploring the Strengths and Weaknesses of AI-Generated Content in Game Development*. Stockholm university.

[31]. Gupta, A. (2023). Are ChatGPT and GPT-4 Good Poker Players?--A Pre-Flop Analysis. *arXiv preprint arXiv:2308.12466*.