**Template para entrega do projeto da disciplina**

**Programação para Dados**

Fase 1

|  |  |
| --- | --- |
| **Nome do estudante** | **Victor Moraes Zucchetti** |
| **Curso** | **Banco de Dados: Ênfase em Data Analytics** |

ESTE É O TEMPLATE PARA PREENCHIMENTO DO PROJETO. PARA PRODUZIR O TRABALHO, ORIENTE-SE PELA DESCRIÇÃO DO ENUNCIADO DA FASE 1 DO PROJETO.

Para demonstrar seu trabalho, devem ser respondidas três perguntas, sobre o conjunto de dados completo, demonstrado a organização proposta:

* Atenção: Além da resposta estar correta, deve discorrer brevemente sobre os resultados obtidos (evite respostas curtas).

**Pergunta 1: Qual o percentual de jogos gratuitos e pagos na plataforma?**

|  |
| --- |
| Resposta:  A partir do dataset completo, foram criadas duas funções para armazenar em uma lista os jogos gratuitos (price = 0) e os jogos pagos (price > 0), retornando o número de jogos na lista. Com isso, foi calculada a porcentagem de gratuitos e pagos através de uma divisão no número de jogos pagos e gratuitos pelo número total de jogos, e o quociente multiplicado por 100.  O dataset apresentou uma grande maioria de jogos pagos. Dos 72934 jogos presentes na plataforma, 60254 eram pagos, o que equivale à 82.61% dos jogos. A plataforma apresentou 12680 dos seus jogos gratuitos, equivalendo à 17.39%. dos jogos totais. |

**Pergunta 2: Qual o ano com o maior número de novos jogos? Em caso de empate, retorne uma lista com os anos empatados.**

|  |
| --- |
| Resposta:  Para obter o(s) ano(s) com maior número de jogos lançados, primeiro foi criada uma função para realizar a contagem de jogos por ano de lançamento e armazenar esses valores por ano em um dicionário. Após, uma segunda função percorre os valores deste dicionário para guardar em uma lista o(s) anos com maior número de jogos e o número de jogos lançados nesse(s) ano(s).  O ano que apresentou o maior número de jogos novos foi 2022, com um total de 13961 jogos lançados. |

**Pergunta 3: Para demonstrar a facilidade de revisão e modificação de uso do módulo desenvolvido, uma pergunta adicional deve ser proposta e respondida por você.**

|  |
| --- |
| Pergunta: **Quais os gêneros com maior média de tempo jogado?**  Resposta:  Para responder esta pergunta, uma função foi utilizada para separar o(s) gênero(s) de cada jogo e salvar valores únicos em uma lista. Uma segunda função adicionou os valores desta lista a um dicionário e percorreu o dataset para encontra jogos que possuíam um gênero específico, para então, fazer a somatória do tempo médio jogado e adicionar ao(s) gênero(s). O total desta soma foi então divido por 60 para converter este valor para horas. Por fim, uma terceira função realizou um ordenamento desses valores de forma decrescente, e armazenou em uma lista os cinco primeiros valores (maiores).  Os 5 gêneros com maior média de tempo de jogo são Indie, com um 61556.25 horas, Action, com 54099.17 horas, Adventure, com 48737.68 horas, Strategy, com 39805.87, e RPG, com 39522.60 horas. |

Como você é um analista de dados dedicado, serão criados testes sobre uma amostra de 20 jogos do cadastro separados aleatoriamente em um arquivo. Responda as mesmas três perguntas, sobre o conjunto de dados da amostra, demonstrado a organização proposta:

* Atenção: Além da resposta estar correta, deve discorrer brevemente sobre os resultados obtidos (evite respostas curtas).

**Pergunta 1: Qual o percentual de jogos gratuitos e pagos na plataforma?**

|  |
| --- |
| Resposta:  Os dados da amostragem realizada com 20 jogos foram contados a mão, com o auxílio do Excel.  Na amostra de dados com 20 jogos, grande parte deste eram pagos, sendo 18 dos 20 (90%), enquanto 2 (10%) são distribuídos de forma gratuita pela plataforma. |

**Pergunta 2: Qual o ano com o maior número de novos jogos? Em caso de empate, retorne uma lista com os anos empatados.**

|  |
| --- |
| Resposta:  Os dados da amostragem realizada com 20 jogos foram contados a mão, com o auxílio do Excel.  A ano com mais jogos lançados nessa amostra de dados foi 2021, com um total de 7 jogos. |

**Pergunta 3: Para demonstrar a facilidade de revisão e modificação de uso do módulo desenvolvido, uma pergunta adicional deve ser proposta e respondida por você.**

|  |
| --- |
| Pergunta: **Quais os gêneros com maior média de tempo jogado?**  Resposta:  Os dados da amostragem realizada com 20 jogos foram contados a mão, com o auxílio do Excel.  Os 5 gêneros com maior média de tempo de jogo são RPG, com 8.28 horas, Indie, com 3.60 horas, Casual, com 3.60 horas, Adventure, com 2.73, e Strategy, com 0.88 horas. |