**Lab 7 Association Rule Learning (10/6/2022)**

**ລະຫັດນັກສຶກສາ: 205Q0010.19**

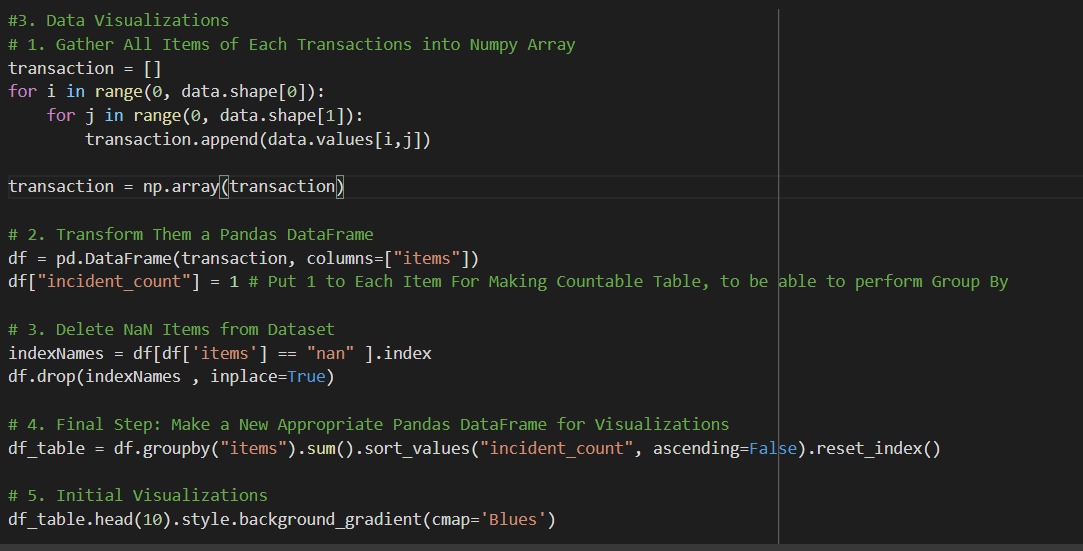
**ຊື່ ແລະ ນາມສະກຸນ:.ທ ນູຊົ່ວ ເຮີ 3CW1**

ຈົ່ງນຳໃຊ້ຄຳສັ່ງຂອງ Python ດ້ວຍ Association Rule Learning.ipynb ເພື່ອຕອບຄຳຖາມຕໍ່ໄປນີ້ໃຫ້ສຳເລັດ:

1.ຈາກຊຸດຂໍ້ມູນ Market\_Basket\_Optimisation.csv. ຈົ່ງບອກຂະໜາດຂອງຊຸດຂໍ້ມູນວ່າມີຈຳນວນແຖວ ແລະ  ຖັນເທົ່າກັບເທົ່າໃດ?

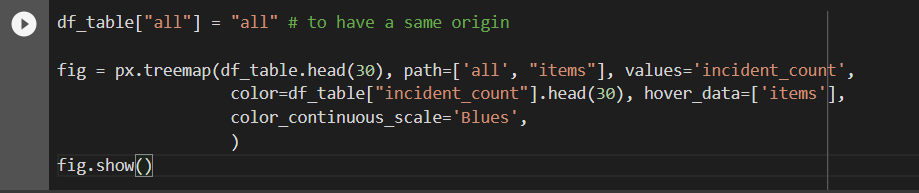
|  |  |
| --- | --- |
| Data.Shape | |
| Lows | 7501 |
| Columns | 20 |

2. ຈົ່ງທຳການສຳຫຼວດຊຸດຂໍ້ມູນດ້ວຍການສະແດງຄ່າຄວາມຖີ່ຂອງ 10 ຜະລິດຕະພັນແຕ່ໃຫຍ່ຫານ້ອຍ.



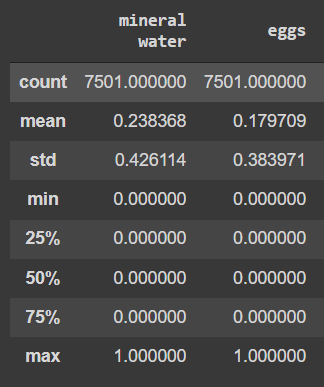


3.ຈົ່ງທຳການສຳຫຼວດຊຸດຂໍ້ມູນດ້ວຍ treemap ແລະ ສະແດງຄ່າຄວາມຖີ່ຂອງ 30 ຜະລິດຕະພັນແຕ່ໃຫຍ່ຫານ້ອຍ



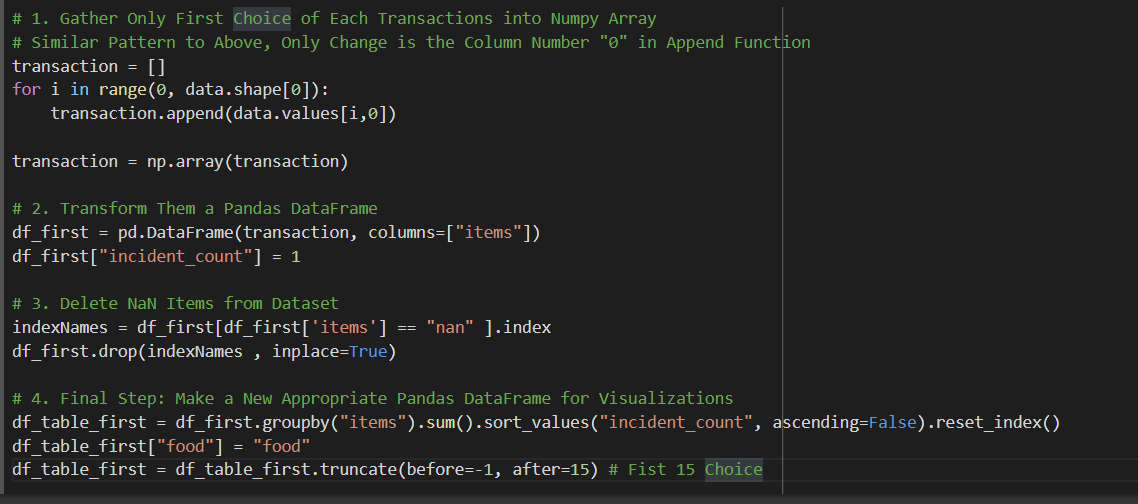


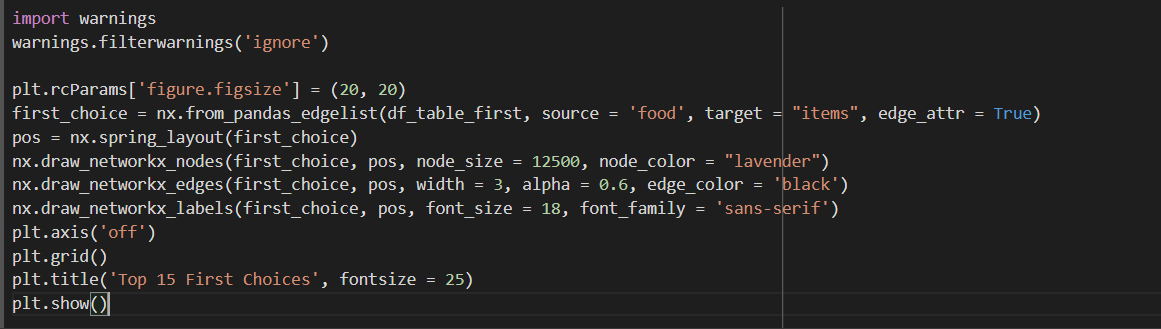
4. ຈົ່ງທຳການສຳຫຼວດຊຸດຂໍ້ມູນດ້ວຍ df\_top20\_multiple\_record\_check.describe(). ຖາມວ່າລະຫວ່າງ ຜະລິດຕະພັນ mineral water ແລະ eggs ຜະລິດຕະພັນໃດມີຄ່າຜັນປ່ຽນມາດຕະຖານນ້ອຍກ່ວາກັນ?

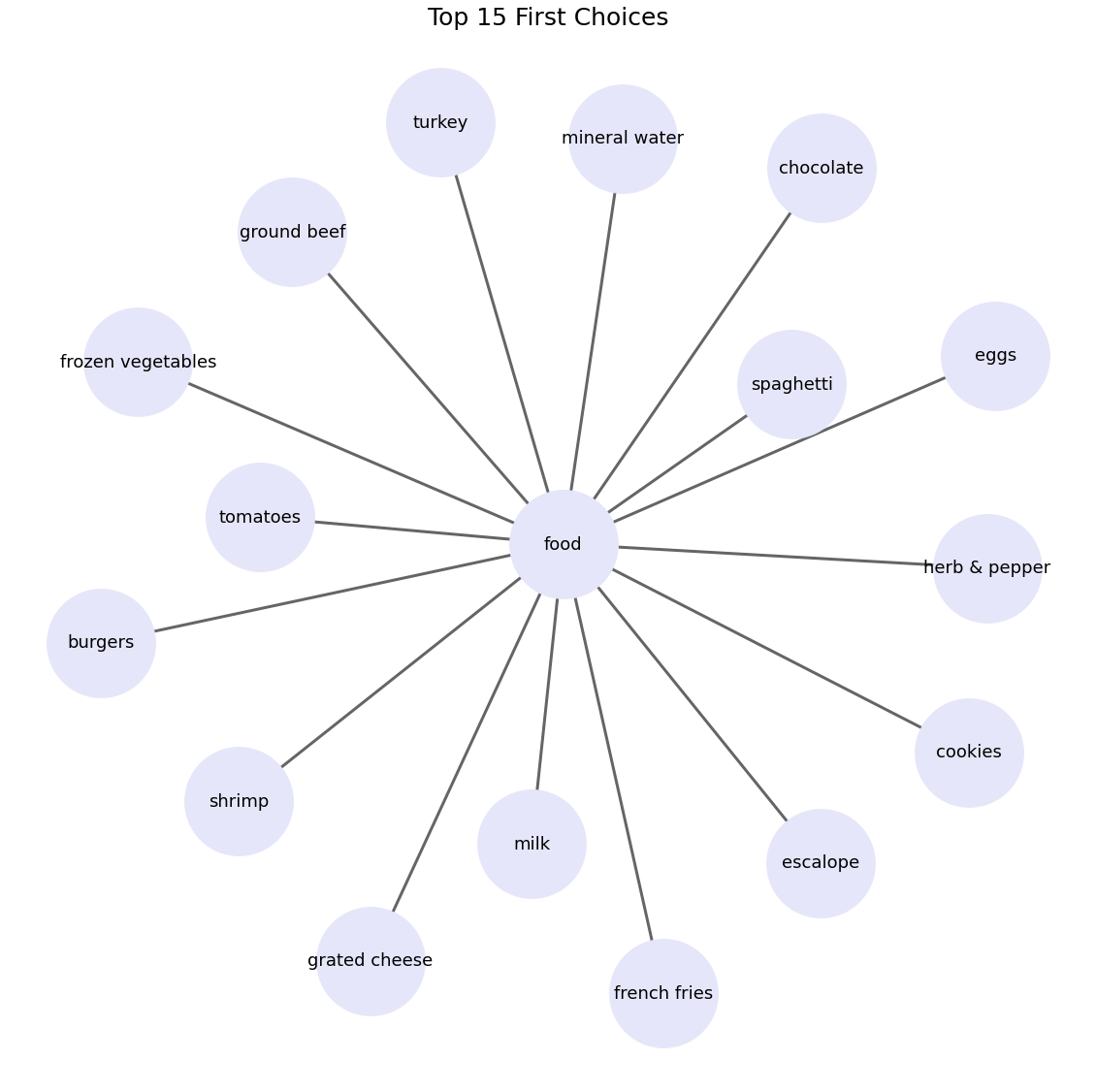


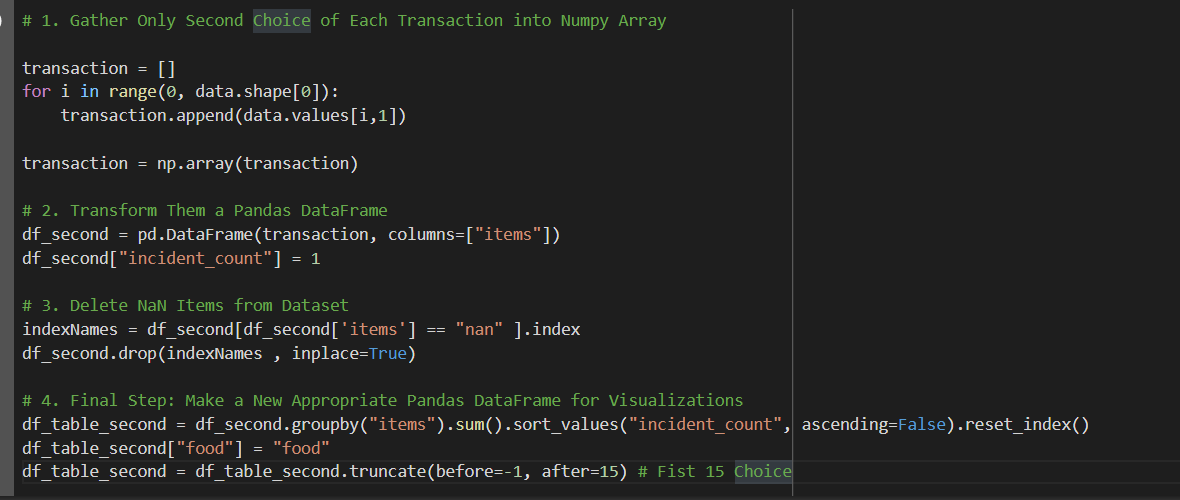
Eggs ມີຄ່າຜັນປ່ຽນມາດຕະຖານນ້ອຍກ່ວາ

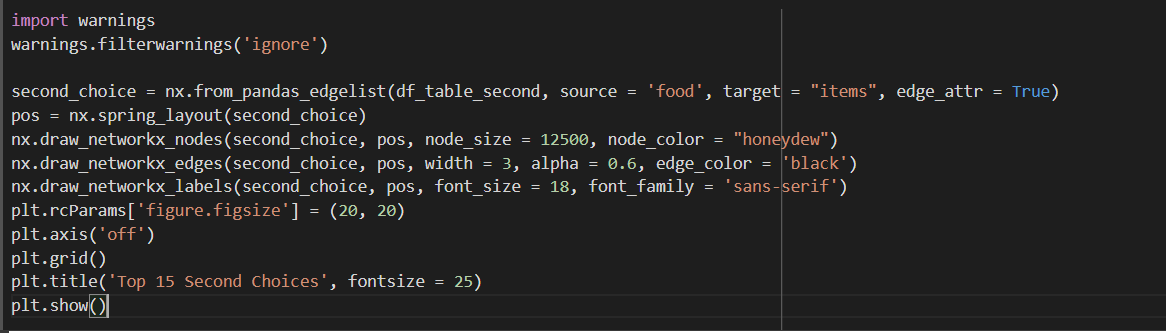
5. ຈົ່ງທຳການສຳຫຼວດຊຸດຂໍ້ມູນດ້ວຍການສະແດງ 15 ຜະລິດຕະພັນໃນການເລືອກຄັ້ງທຳອິດ (Top 15 First  Choices) ແລະ 15 ຜະລິດຕະພັນໃນການເລືອກຄັ້ງທີ 2 (Top 15 Second Choices) ຂອງລູກຄ້າ.

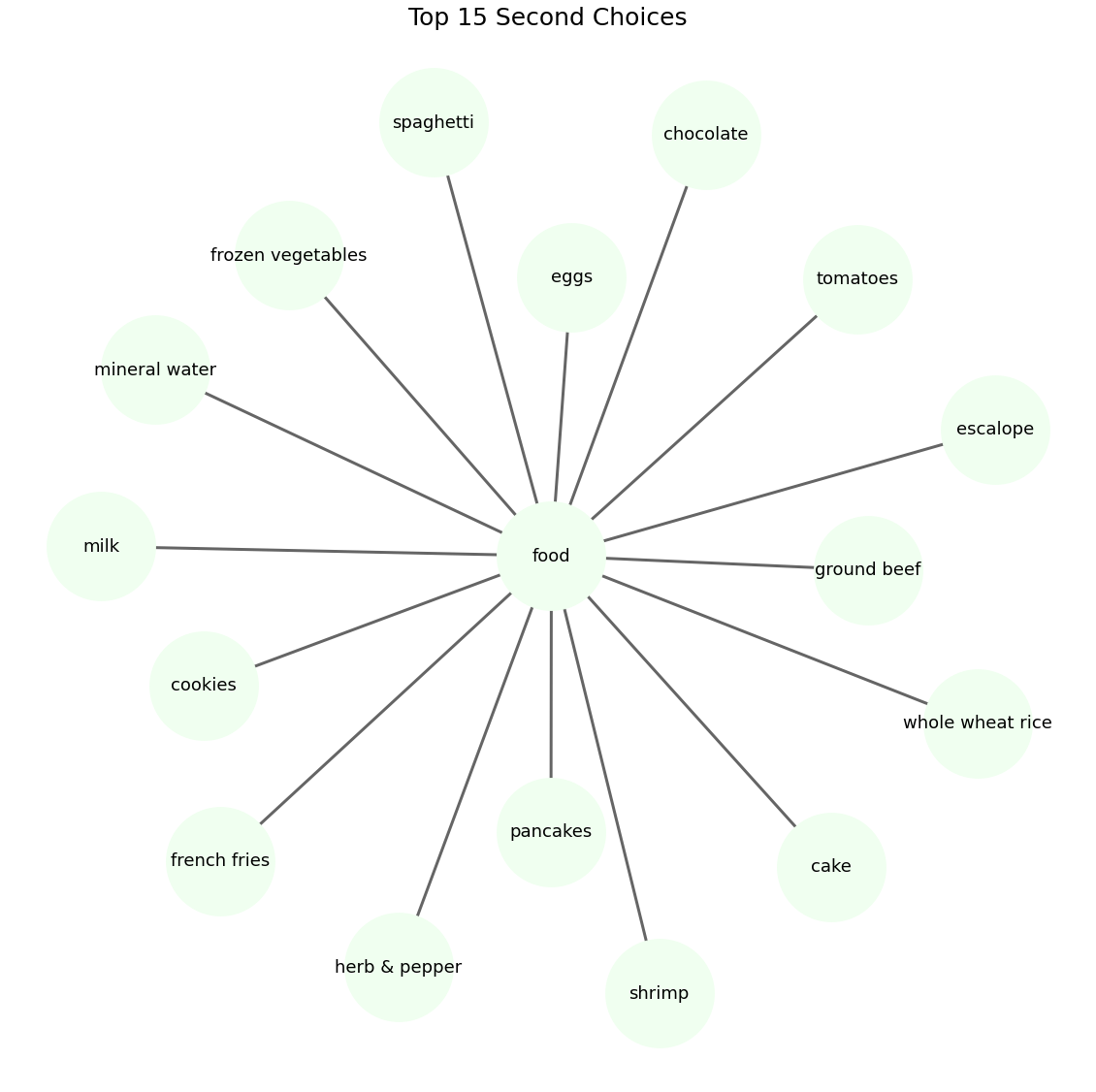




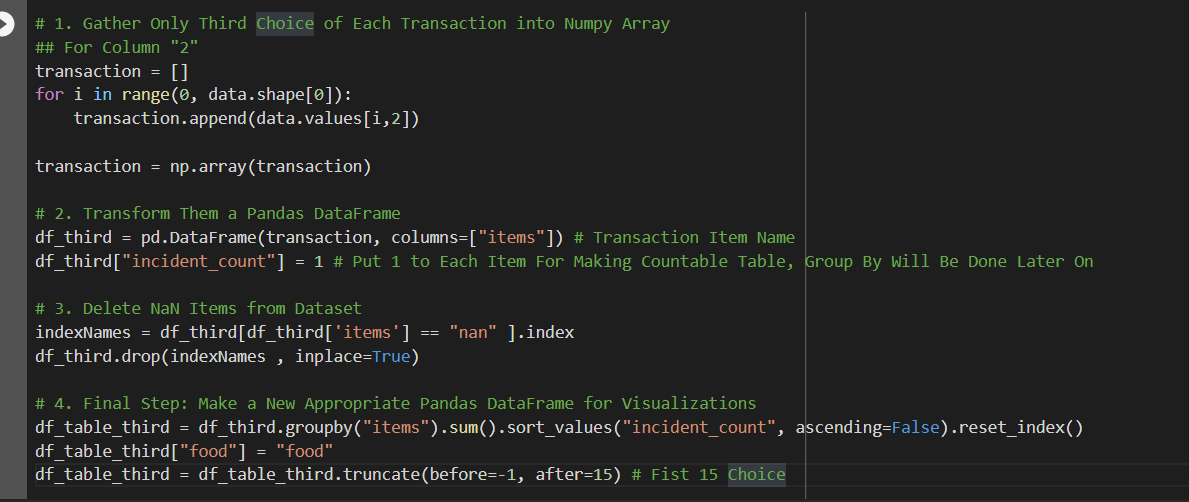


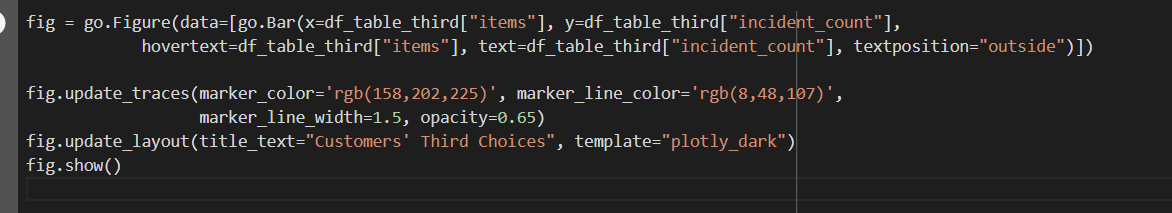


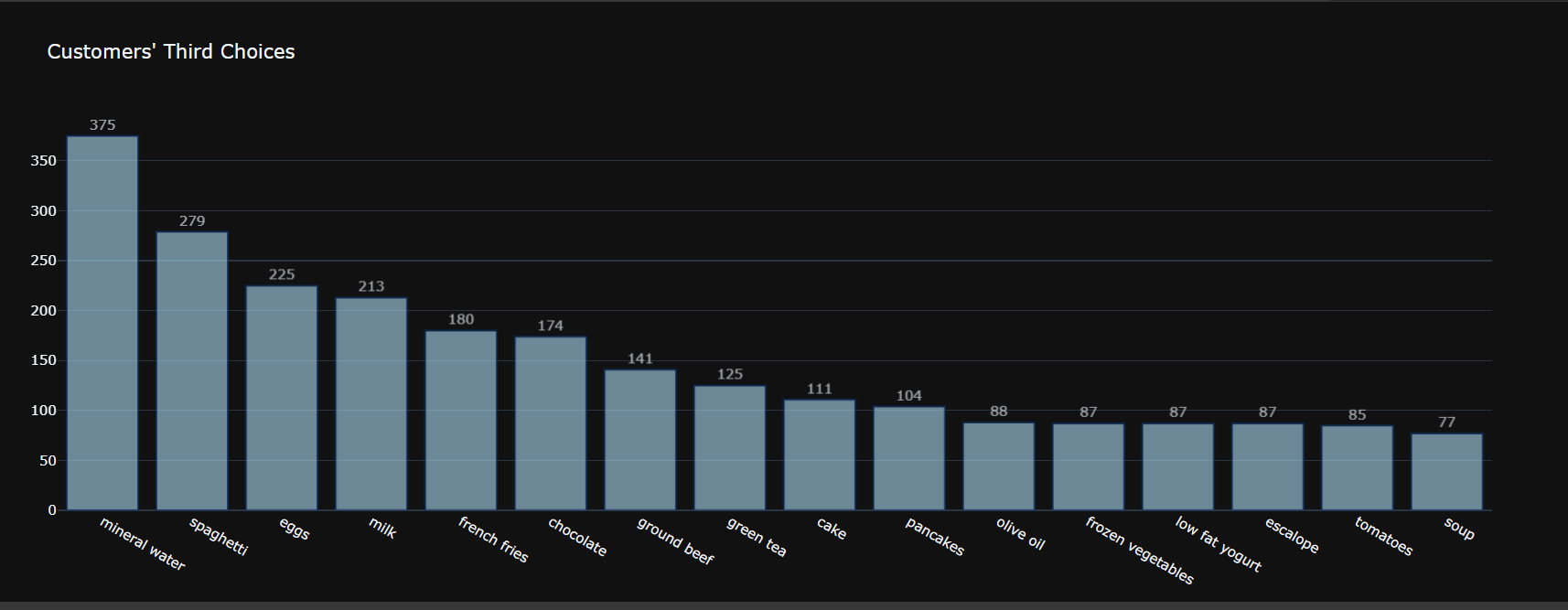




6. ຈົ່ງທຳການສຳຫຼວດຊຸດຂໍ້ມູນດ້ວຍການສະແດງຄ່າຄວາມຖີ່ຂອງຜະລິດຕະພັນໃນການເລືອກຄັ້ງທີ 3 ຂອງລູກຄ້າດ້ວຍ  (Customers' Third Choices).







7. ຖາມວ່າກ່ອນການກະກຽມຂໍ້ມູນ (data preprocessing) ຊຸດຂໍ້ມູນມີ memory usage ເທົ່າໃດ?

memory usage: 1.1 MB

8. ຖາມວ່າຫຼັງການກະກຽມຂໍ້ມູນ (data preprocessing) ຊຸດຂໍ້ມູນມີ memory usage ເທົ່າໃດ?

memory usage: 886.5 KB

9. ຖາມວ່າຫຼັງການຂັດເລືອກເອົາ 50 ຜະລິດຕະພັນໃນຊຸດຂໍ້ມູນມາດຳເນີນການປະມວນຜົນ, memory usage ເທົ່າໃດ (ຄ່າຂໍ້ມູນເປັນ boolean)?

memory usage: 366.4 KB

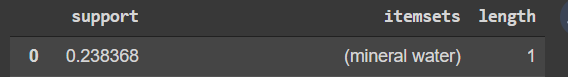
10. ຖາມວ່າຫຼັງການຂັດເລືອກເອົາ 50 ຜະລິດຕະພັນໃນຊຸດຂໍ້ມູນມາດຳເນີນການປະມວນຜົນ, memory usage ເທົ່າໃດ (ຄ່າຂໍ້ມູນເປັນ int64)?

memory usage: 2.9 MB

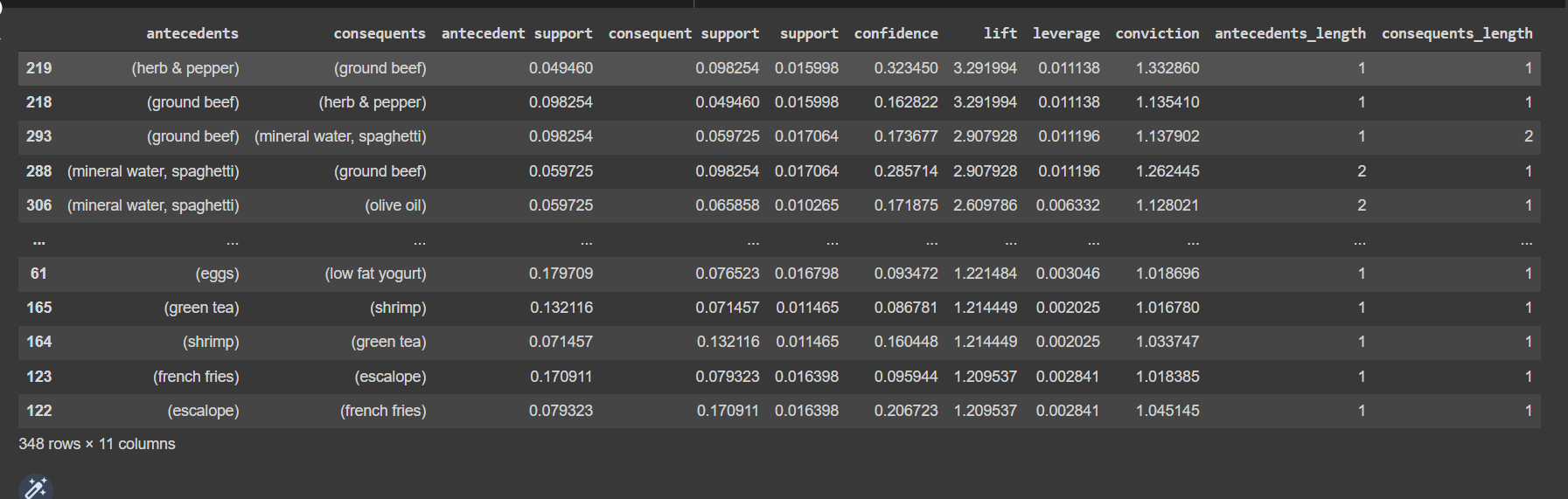
11. ໃນການວິເຄາະຄວາມສຳພັນຂອງຂໍ້ມູນດ້ວຍ Association Rule Learning ຄັ້ງນີ້. ຖາມວ່າເພີ່ນນຳໃຊ້  Algorithm ຕົວໃດ?

ຕອບ: Apriori Algorithm

12. ຖາມວ່າຜະລິດຕະພັນ mineral water ມີຄ່າ support ເທົ່າໃດ?



13. ເມື່ອ antecedents ເປັນ (mineral water, spaghetti) ເວລານັ້ນ consequents ເປັນ olive oil. ຖາມ ວ່າ confidence ແລະ lift ມີຄ່າເທົ່າໃດ?



* Confidence 0.171875
* Lift 2.609786

14. ຖາມວ່າຄວາມສຳພັນຂອງກຸ່ມຜະລິດຕະພັນໃດມີຄ່າ confidence ສູງກ່ວາໝູ່?

