**Lab 6.1 Hierarchical Clustering**

**(27/5/2022)**

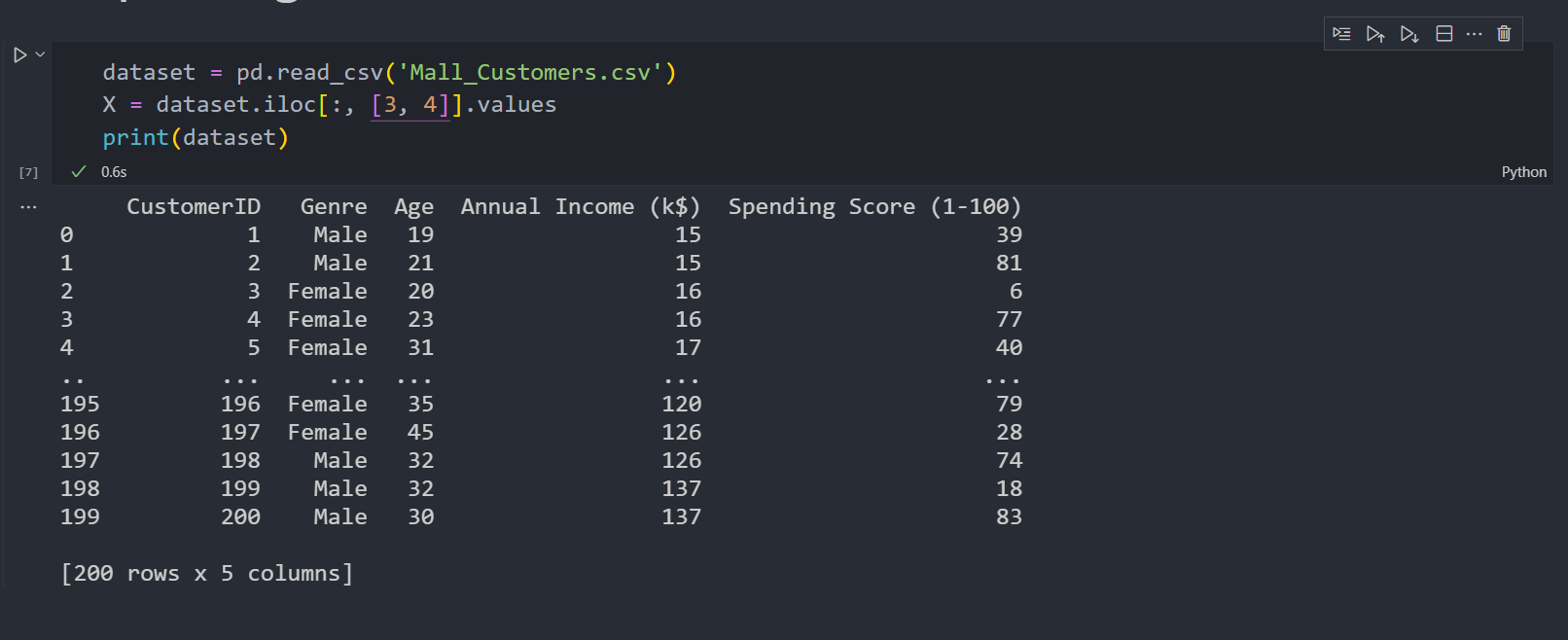
**ລະຫັດນັກສຶກສາ: 205N0083.19**

**ຊື່ ແລະ ນາມສະກຸນ: ທ ເຊັງວ່າງ ບຼົ່ງໄມ 3CW1**

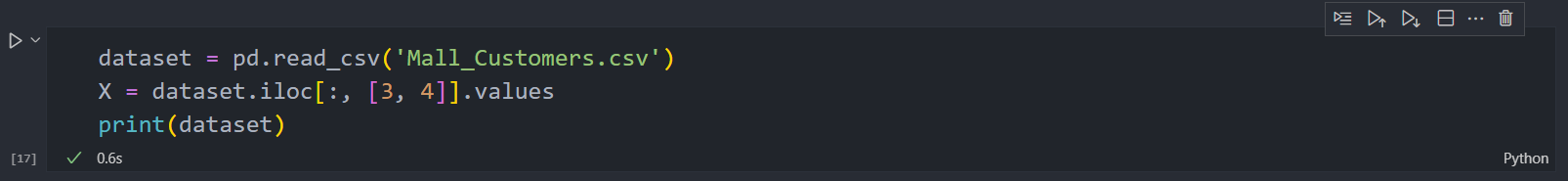
ຈົ່ງຕອບຄຳຖາມຕໍ່ໄປນີ້ໃຫ້ສຳເລັດດ້ວຍການນຳໃຊ້ຄຳສັ່ງຂອງ Python:

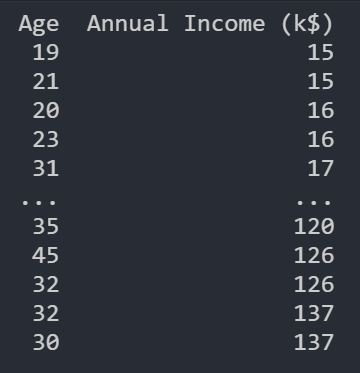
ພາກທີ 1

* 1. ຈາກຊຸດຂໍ້ມູນMall\_Customers.csvຈົ່ງເຕີ່ມຂໍ້ມູນໃນຕາຕະລາງໃຫ້ສຳເລັດ.

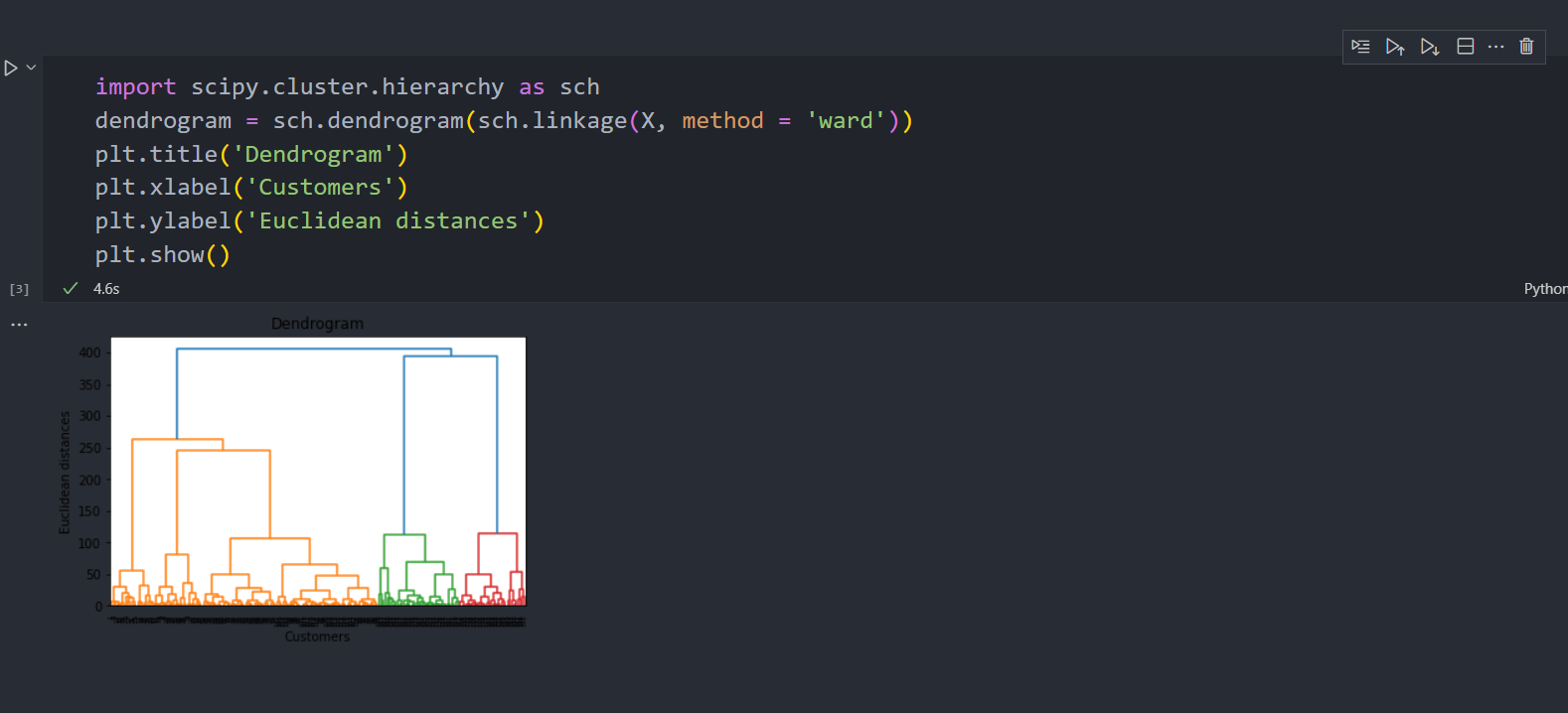


* 1. ຈົ່ງທຳການເລືອກຄຸນລັກສະນະ (Features): Age ແລະ Annual Income (k$)

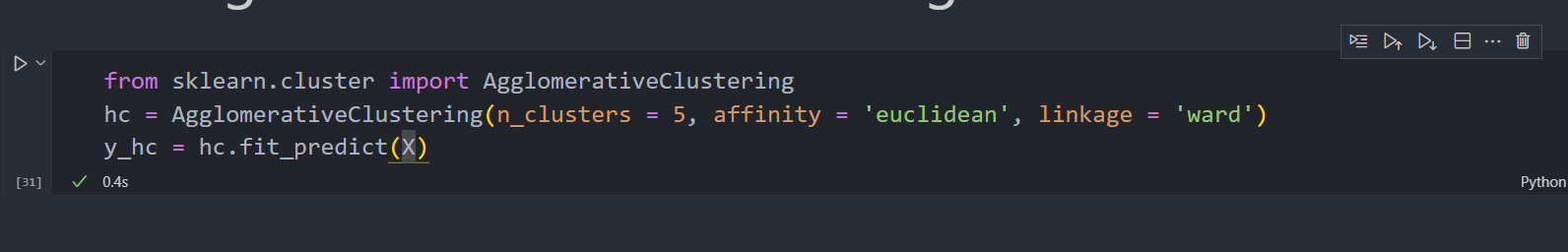




* 1. ຈົ່ງຊອກຫາຈຳນວນກຸ່ມເພື່ອແບ່ງກຸ່ມຊຸດຂໍ້ມູນໃຫ້ເມາະສົມດ້ວຍDendrogram



1.4. ຈົ່ງທຳການຝຶກຂໍ້ມູນໃນຂໍ້ທີ 1.2 ເພື່ອແບ່ງຕາມຈຳນວນກຸ່ມໃນຂໍ້ທີ 1.3 ດ້ວຍແບບຈຳຮອງ (Algorithm) AgglomerativeClustering.



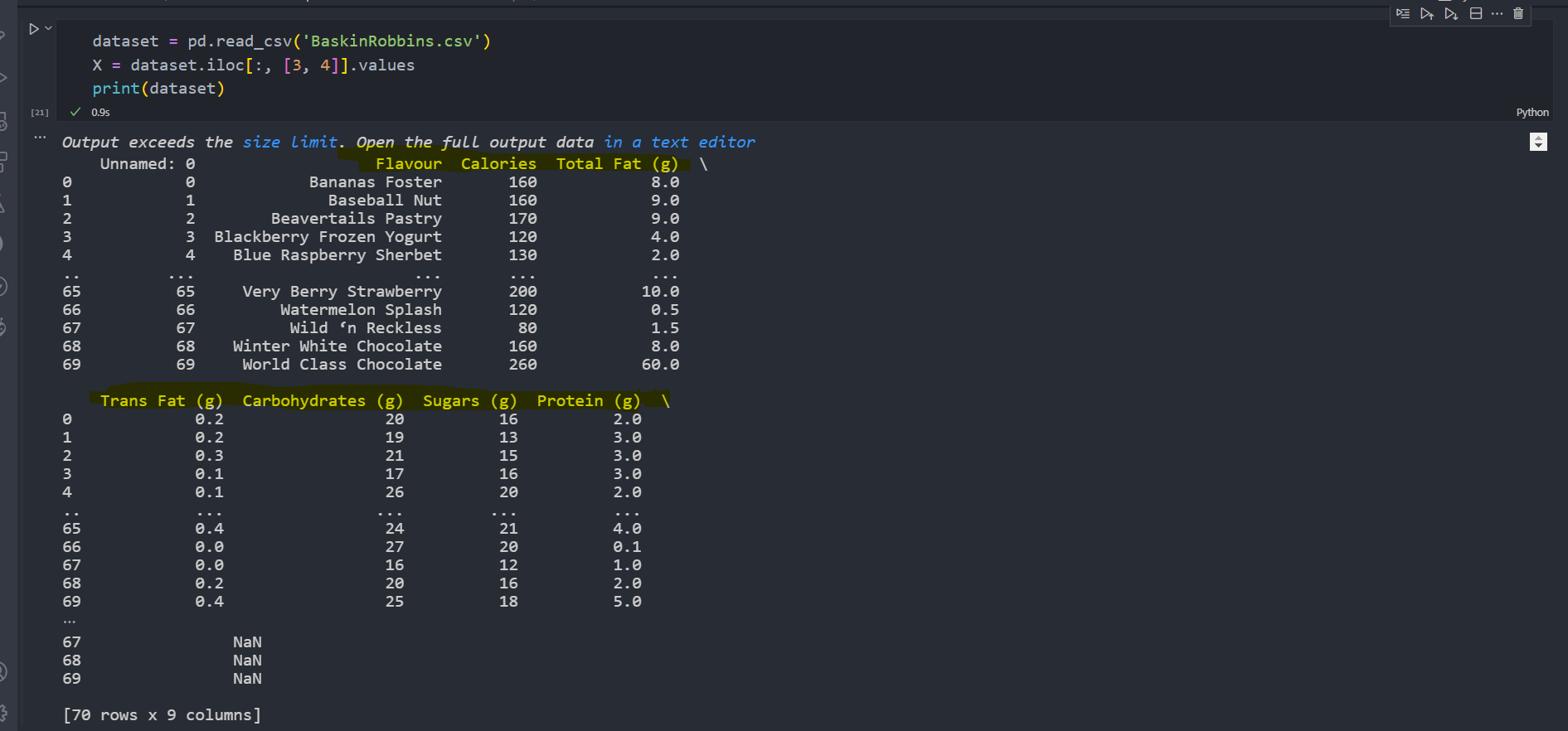
1.5. ຈົ່ງທຳການສຳຫຼວດ (visulaizing) ກຸ່ມຂໍ້ມູນທີ່ຖືກແບ່ງດ້ວຍscatter plot ພ້ອມອະທິບາຍຜົນ.



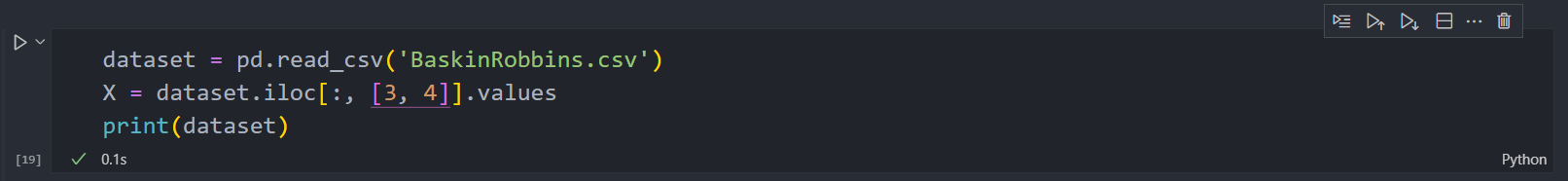
* **Cluster1 (ສີເເດງ): ລາຍຮັບສູງ ແລະ ຄະແນນການໃຊ້ຈ່າຍຕໍ່າ**
* **Cluster2 (ສີຟ້າ): ລາຍໄດ້ປົກກະຕິ ແລະ ຄະແນນການໃຊ້ຈ່າຍປົກກະຕິ**
* **Cluster3 (ສີຂຽວ): ລາຍຮັບສູງ ແລະ ຄະແນນການໃຊ້ຈ່າຍສູງ**
* **Cluster4 (ສີຟ້າຂຽວ): ລາຍຮັບຕໍ່າ ແລະ ຄະແນນການໃຊ້ຈ່າຍສູງ**
* **Cluster5 (ສີມ່ວງແດງ): ລາຍຮັບຕໍ່າ ແລະຄະແນນການໃຊ້ຈ່າຍຕໍ່າ**

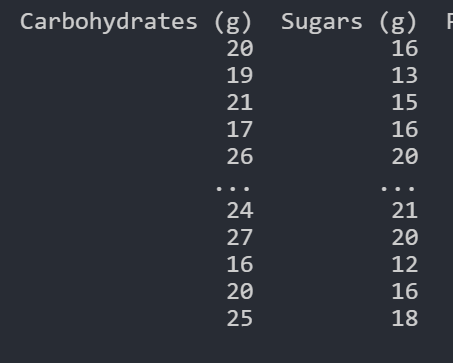
ພາກທີ 2

2.1ຈາກຊຸດຂໍ້ມູນBaskinRobbins.csvຈົ່ງເຕີ່ມຂໍ້ມູນໃນຕາຕະລາງໃຫ້ສຳເລັດ.



2.2. ຈົ່ງທຳການເລືອກຄຸນລັກສະນະ (Features): Carbohydrates (g) ແລະ Sugars (g)

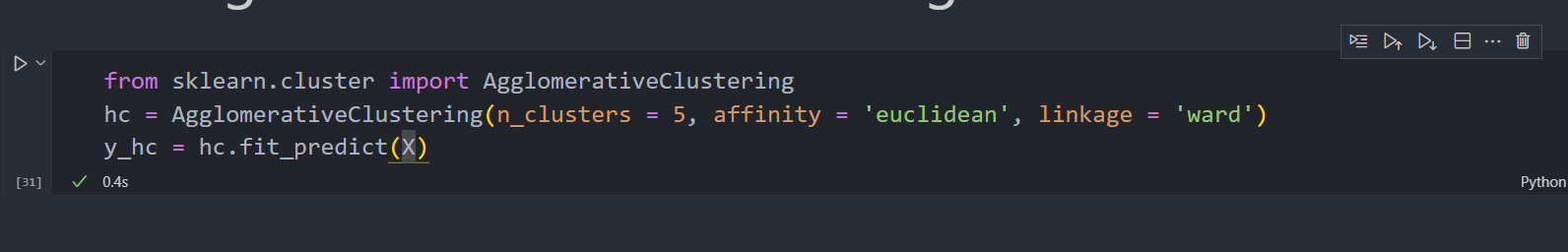




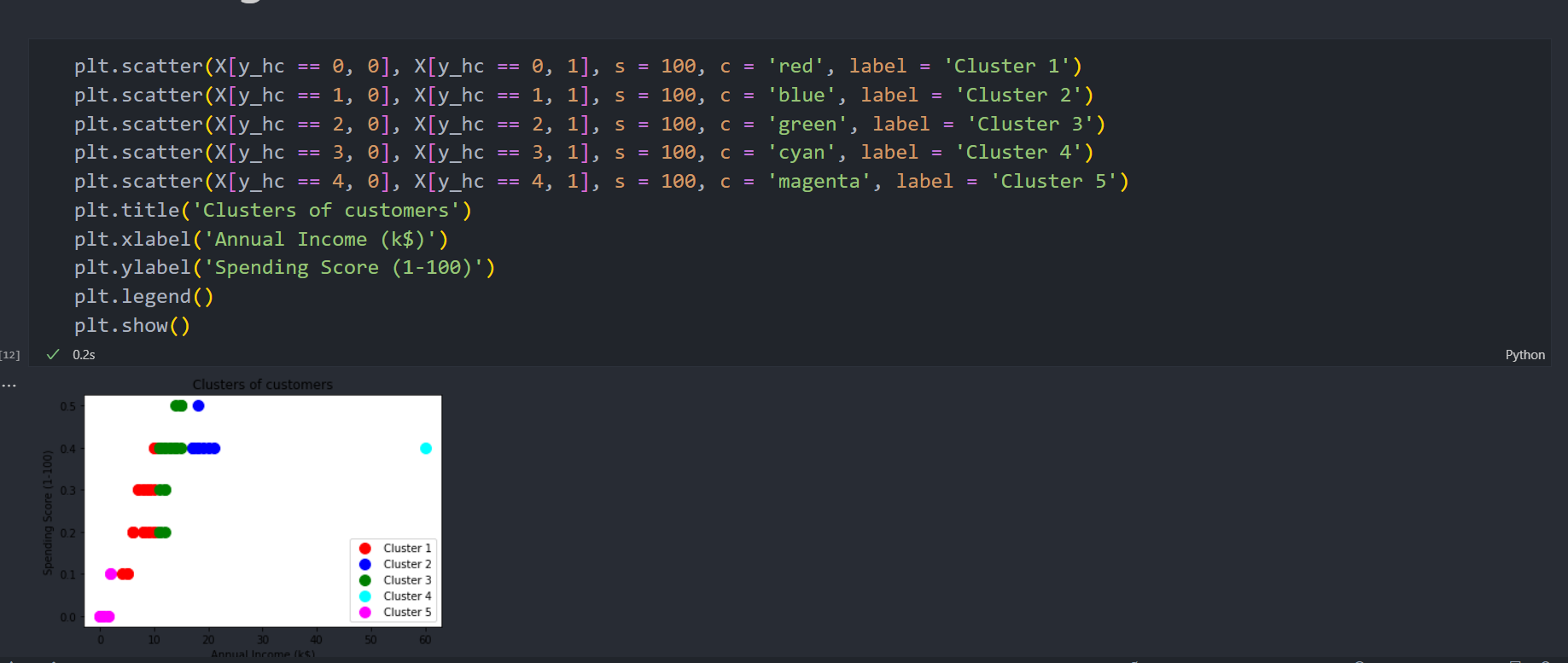
2.3. ຈົ່ງຊອກຫາຈຳນວນກຸ່ມເພື່ອແບ່ງກຸ່ມຊຸດຂໍ້ມູນໃຫ້ເມາະສົມດ້ວຍDendrogram



2.4. ຈົ່ງທຳການຝຶກຂໍ້ມູນໃນຂໍ້ທີ 2.2 ເພື່ອແບ່ງຕາມຈຳນວນກຸ່ມໃນຂໍ້ທີ 2.3 ດ້ວຍແບບຈຳຮອງ (Algorithm) AgglomerativeClustering.



2.5. ຈົ່ງທຳການສຳຫຼວດ (visulaizing) ກຸ່ມຂໍ້ມູນທີ່ຖືກແບ່ງດ້ວຍ scatter plot ພ້ອມອະທິບາຍຜົນ.

* 
* **Cluster1 (ສີເເດງ): ລາຍຮັບຕໍ່****າ ແລະ ຄະແນນການໃຊ້ຈ່າຍບາງເທື່ອກໍ່ສູງບາງເທື່ອກໍ່ຕໍ່າ**
* **Cluster2 (ສີຟ້າ): ລາຍໄດ້ປົກກະຕິ ແລະ ຄະແນນການໃຊ້ຈ່າຍສູງ**
* **Cluster3 (ສີຂຽວ): ລາຍຮັບຕໍ່າ ແລະ ຄະແນນການໃຊ້ຈ່າຍສູງ**
* **Cluster4 (ສີຟ້າຂຽວ): ລາຍຮັບຕໍ່າ ແລະ ຄະແນນການໃຊ້ຈ່າຍສູງ**
* **Cluster5 (ສີມ່ວງແດງ): ລາຍຮັບສູງ ແລະຄະແນນການໃຊ້ຈ່າຍສູງ**