**Assistance from**

|  |  |
| --- | --- |
| * 1-on-1 tutor Geronimo Perez | * Office hours with TA Vicky Liao |
| * Tuition session with Michael Wenz | * ChatGPT |

**YouTube Videos & Cheat Sheets**

1. Links from Mark Jenson:

* [Recommended Book: High Performance Spark](http://url1573.bootcampspot.com/ls/click?upn=zZ9isbxdmSyhQs8zZnactHDuFLBWlIu4srL8VgZeffFfFYQu2mxD7HqeyWuRWJksuh6ZjIM3XoJKmuOxWTtv0REW6OiQA8yOL5MWwWzgrrJbJgiNj0G3J5nkhlbRUalNn2o6_yNRBD6oYCC4Q2oCzDO9te2UwjKXZOk8UDUlqB3f9sNfkM3vzrozaHEoApFR-2FmvOeUwr9JOi1-2BeQpo6B95Re1bKQBmXkpGsfInB9vOx8Z6lbDRLb9W393JTDM5v0psKQiUY4WIcAow71epSQc86VC1FxUt3vR7II0QjG6EvOgjS5rhIKUyMnpcVNdHlgb9iFfG0B468rwVdKRzYUDTYBtVg-3D-3D)
* [Apache Spark Performance Tuning](http://url1573.bootcampspot.com/ls/click?upn=zZ9isbxdmSyhQs8zZnactDcumMzjWe76EsUAjD5VHeG5v9eak3jdG3B76EGGoWvvhh6lv9By6Uty1XK-2B1yMffnFRFS0RTIhuByENkI-2F0BoM-3D9N9Y_yNRBD6oYCC4Q2oCzDO9te2UwjKXZOk8UDUlqB3f9sNfkM3vzrozaHEoApFR-2FmvOeSLim8jcPyNt3AMc7SkfMzsjDJvGs3frFtu6SWshPonaDUGPKgqteIlMSf1GEC2CLjRK7lseT3s7WLQW3meaUrB9SYSYQB5W4RQUp6EFgtAk338pTgHMY2hw4zG9ElkWOt4H80Z0Lv-2FnDO5al4tAyeg-3D-3D)
* [Best practices to to optimizing query performance](http://url1573.bootcampspot.com/ls/click?upn=zZ9isbxdmSyhQs8zZnactMGn0VlGiw0yKsKCKwkJYBpe0q6rC3BubIOWVlfLrT3RhL6p-2B0YmCIXoZERTPj0LBiKvCRza-2BonwwVgeoglbwwKI6gXr-2FrjYost3gScenoaBzB8G_yNRBD6oYCC4Q2oCzDO9te2UwjKXZOk8UDUlqB3f9sNfkM3vzrozaHEoApFR-2FmvOeiP47hMcwZg2cz3G00hqZL8W-2BHa1a-2FM8MaoDSJ1Rhl0TYb7BZS7es2zqvNzOmboX8voxvuvxmyotTOzZTWZYDo2mkh6G2KIOUW1hEQ5soKtFCw2ROKJ00bbtNRso-2FF-2B2-2Bs4Nb1wvOkpF1yWBXvPd3DA-3D-3D)
* [Parquet and Partitioning Best Practices](http://url1573.bootcampspot.com/ls/click?upn=zZ9isbxdmSyhQs8zZnactG5FyfyCQd-2BNsrjFbGJ-2B-2FmRb1KKQos6pxhA3L-2FC1MoHZeVNSSywdI-2B0EbY-2Fxs2rz-2B9l1wzz0xVFfY3wE2qgmVwY-3DcoUT_yNRBD6oYCC4Q2oCzDO9te2UwjKXZOk8UDUlqB3f9sNfkM3vzrozaHEoApFR-2FmvOeA6Bqo1lgPMmRK4Ifb-2FJzmTt5AWkuK8-2BUiNL29Dmm99K3KrpMLfGL62FVD1YtLDTp0TczX-2Fmq3oL6ylFLHSwlBpLvnTANVCOJYsBHEX0TJWXjvyAJUpjtTpAEPCogJtPxLnRLp9MFmrpsk6fnfr-2Bfsg-3D-3D)

1. PySpark\_SQL\_Cheat\_Sheet: <https://git.bootcampcontent.com/University-of-Adelaide/UADEL-VIRT-DATA-PT-06-2023-U-LOLC/-/blob/main/01-Lesson-Plans/22-Big-Data/2/PySpark_SQL_Cheat_Sheet.pdf>
2. Hadoop In 5 Minutes: <https://www.youtube.com/watch?v=aReuLtY0YMI>
3. What exactly is Apache Spark? | Big Data Tools: <https://www.youtube.com/watch?v=ymtq8yjmD9I>
4. What Is Apache Spark? <https://www.youtube.com/watch?v=VZ7EHLdrVo0>

<https://www.youtube.com/watch?v=SxAxAhn-BDU>

1. Learn Apache Spark in 10 Minutes | Step by Step Guide: <https://www.youtube.com/watch?v=v_uodKAywXA>

**In-Class Activities**

Module 22 Class 1 Activities 07

Module 22 Class 2 Activities 02, 08

Module 22 Class 3 Activities 01, 02, 04, 06