

Lecture #	Date	Day	Assignment	Papers to Read and Write a Report On	Quiz	Category	Topics	Reference Videos
1	18-Jan	Tuesday	Setting up colab and AWS and testing the system works	Towards felicitous decision making: An overview on challenges and trends of Big Data + Big Data Systems: A Software Engineering Perspective		Introduction	Overview of the class+ Understanding Big Data	https://youtu.be/0uc9zB-FmVk
2	20-Jan	Thursday					Applications Of Big Data: Wildfire + Transit Analysis + Emergency Response	
3	25-Jan	Tuesday	Assignment 1: sql	Anatomy of a Database System		Storage	Overview of databases and their anatomy+ACID Properties	https://www.youtube.com/watch?v=5Pia4UFuMKo
4	27-Jan	Thursday		Access Path Selection and Joins+ Optimistic Concurrency			Relational Databases : SQL and Joins and access	https://youtu.be/HXV3zeQKqGY
5	1-Feb	Tuesday					Relational Databases: Indexes, Locks and Concurrency -00	https://youtu.be/ihXf8p_-UwU
6	3-Feb	Thursday					Distribution/Partitioning: Fault Tolerance+ CAP Theorem	https://www.youtube.com/watch?v=R_Fxz14tr2M
7	8-Feb	Tuesday	Column vs Row Oriented Databases				https://www.youtube.com/watch?v=Vw1fCeD06YI	
8	10-Feb	Thursday	Assignment 2: MongoDB	columnstoresfntdbs (chapters 1, 2 and 3)			Introducing NoSQL Databases: Need and Differences/ MongoDB	https://www.youtube.com/watch?v=QwevGzVu_zk
9	15-Feb	Tuesday		NOSQL Paper			BigTable	https://www.youtube.com/watch?v=KaRbKdMInuc
10	17-Feb	Thursday					BigTable and DynamoDB	DynamoDB: https://www.youtube.com/watch?v=OfZgHXsYqNE
11	22-Feb	Tuesday			InfluxDB		https://www.youtube.com/watch?v=OoCsY8odmpM	
12	24-Feb	Thursday	Assignment 3: Influx and Pandas	Graph DataBase Paper	Quiz 1	Graph Databases	https://www.youtube.com/watch?v=GekQqFZm7mA	
13	1-Mar	Tuesday		Fault Tolerance Paper			Overview of Pandas and Parquet	Pandas: https://www.youtube.com/watch?v=vmEHCJofslg
14	3-Mar	Thursday					Project Overview and Discussions	Parquet: https://www.youtube.com/watch?v=pFS-FScophU
							Map Reduce	https://www.youtube.com/watch?v=b-lvmXoO0bU
Spring Break								
15	15-Mar	Tuesday	Assignment 4: Hadoop and MapReduce	Orchestrating Big Data	Quiz 2	Batch Processing	Map Reduce	
16	17-Mar	Thursday		map-reduce-bookchapter.pdf + MapReduce: Simplified Data Processing on Large Clusters			Spark	https://www.youtube.com/watch?v=IQfG0faDrzE
17	22-Mar	Tuesday					Dask	https://www.youtube.com/watch?v=EybGGLbLipI
18	24-Mar	Thursday			Assignment 5: Spark and comparison with DASK	ComparisonOfApproachesToLargeScaleDataAnalysis.pdf		Streaming
19	29-Mar	Tuesday	Kafka and Pulsar and ZooKeeper	Kafka: https://www.youtube.com/watch?v=JaIUUBKdcAQ				
20	31-Mar	Thursday	Pulsar	Pulsar: https://www.youtube.com/watch?v=IaVPIY1PRdo				
21	5-Apr	Tuesday	SparkStreaming	https://www.youtube.com/watch?v=UuRhEmqqhRM				
22	7-Apr	Thursday	Storm	https://www.youtube.com/watch?v=0mIEUibjtzk				
23	12-Apr	Tuesday	Assignment 6: Kafka and Spark Streaming	Kafka: a Distributed Messaging System for Log Processing	Quiz 3	Analytics and Real-world Use Cases	Apache Flink	https://www.youtube.com/watch?v=G77m6Ou_kFA
24	14-Apr	Thursday		ZooKeeper: Wait-free coordination for Internet-scale systems			Clustering	https://www.youtube.com/watch?v=eg8DJYwdMyg
25	19-Apr	Tuesday					Pagerank	https://www.youtube.com/watch?v=meonLcN7LD4
26	21-Apr	Thursday			Anomaly Detection		https://www.youtube.com/watch?v=12Xq9OLdQwQ	
27	26-Apr	Tuesday	Assignment 7 (Bonus): Visualization	Efficient Data Management for Intelligent Urban Mobility Systems		Project Presentations	GeoInformational Systems Processing (Traffic and Incident Data)	None
28	28-Apr	Thursday					Project Presentations 1	None
							Project Presetation 2	None