

Lecture #	Date	Day	Assignment	Papers to Read and Write a Report On	Quiz	Category	Topics	Reference Videos
1	26-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	28-Jan	Thursday					Applications Of Big Data: Wildfire + Transit Analysis + Emergency Response	
3	2-Feb	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	4-Feb	Thursday					Relational Databases : SQL and Joins and access	https://youtu.be/HXV3zeQKqGY
5	9-Feb	Tuesday		Access Path Selection and Joins+ Optimistic Concurrency			Relational Databases: Indexes, Locks and Concurrency -00	https://youtu.be/ihXf8p_-UwU
6	11-Feb	Thursday	Assignment 2: MongoDB	columnstoresfntdbs (chapters 1, 2 and 3)			Distribution/Partitioning: Fault Tolerance+ CAP Theorem	https://www.youtube.com/watch?v=R_Fxz14tr2M
7	16-Feb	Tuesday					Column vs Row Oriented Databases	https://www.youtube.com/watch?v=Vw1fCeD06YI
8	18-Feb	Thursday		NOSQL Paper			Introducing NoSQL Databases: Need and Differences/ MongoDB	https://www.youtube.com/watch?v=QwevGzVu_zk
9	23-Feb	Tuesday					BigTable and DynamoDB	BigTable: https://www.youtube.com/watch?v=KaRbKdMInuc DynamoDB: https://www.youtube.com/watch?v=OfZgHXsYqNE
10	25-Feb	Thursday					InfluxDB	https://www.youtube.com/watch?v=OoCsY8odmpM
11	2-Mar	Tuesday	Assignment 3: Influx and Pandas	Graph DataBase Paper	Quiz 1		Graph Databases	https://www.youtube.com/watch?v=GekQqFZm7mA
12	4-Mar	Thursday			Batch Processing		Overview of Pandas and Parquet	Pandas: https://www.youtube.com/watch?v=vmEHCJofslg Parquet: https://www.youtube.com/watch?v=pFS-FScophU
13	9-Mar	Tuesday		Fault Tolerance Paper		Project Overview and Discussions		
14	11-Mar	Thursday	Orchestrating Big Data			Map Reduce		https://www.youtube.com/watch?v=b-lvmXoO0bU
15	16-Mar	Tuesday		Map Reduce				
16	18-Mar	Thursday		Spark		https://www.youtube.com/watch?v=lQfG0faDrzE		
17	23-Mar	Tuesday	map-reduce-bookchapter.pdf + MapReduce: Simplified Data Processing on Large Clusters	Quiz 2	Dask	https://www.youtube.com/watch?v=EybGGGLbLipI		
18	25-Mar	Thursday		Assignment 5: Spark and comparison with DASK	Streaming	Overview (Dataflow Model) + Kafka and Pulsar and ZooKeeper		
19	30-Mar	Tuesday	ComparisonOfApproachesToLargeScaleDataAnalysis.pdf			Kafka and Pulsar and ZooKeeper	Kafka: https://www.youtube.com/watch?v=JaIUUBKdcAQ Pulsar: https://www.youtube.com/watch?v=lAVPIY1PRdo	
20	1-Apr	Thursday				SparkStreaming	https://www.youtube.com/watch?v=UuRhEmqghRM	
21	6-Apr	Tuesday				Storm	https://www.youtube.com/watch?v=0mlEUibjtzk	
22	8-Apr	Thursday	spark.pdf	Quiz 3		Apache Flink	https://www.youtube.com/watch?v=G77m6Ou_kFA	
23	13-Apr	Tuesday	Assignment 6: Kafka and Spark Streaming	Kafka: a Distributed Messaging System for Log Processing	Analytics and Real-world Use Cases	Clustering	https://www.youtube.com/watch?v=eg8DJYwdMyg	
24	15-Apr	Thursday				Pagerank	https://www.youtube.com/watch?v=meonLcN7LD4	
25	20-Apr	Tuesday		ZooKeeper: Wait-free coordination for Internet-scale systems		Anomaly Detection	https://www.youtube.com/watch?v=12Xq9OLdQwQ	
26	22-Apr	Thursday	GeoInformational Systems Processing (Traffic and Incident Data)			None		
27	27-Apr	Tuesday	Assignment 7 (Bonus): Visualization	Efficient Data Management for Intelligent Urban Mobility Systems		Project Presentations	Project Presentations 1	None
28	29-Apr	Thursday					Project Presetation 2	None