100 DAY SELF STUDY MASTER PLAN TO BECOME AZURE DATA ENGINEER(ARUN KUMAR)

Day	Python	SQL	Azure
Day1		Database, Datawarehouse and Data Lake	What is Cloud?
Day1		Type of databases	What are the different major Cloud Platforms in the market?
Day1		RDBMS	What are Iaas,Paas and Saas?
Day1		Logical Schema Designing	Pricing Model in Azure
Day1		Data modeling	
Day2		Star Schema	Create Account in Azure Portal(You will need ceradit card.Rs 2 will be deducted for authentication purpose and that will be returned back.)
Day2		Fact & Dimension table	Azure Portal Overview
Day2		Normalization	Overview of Different Services relevant to Data Engineering in Azure
Day2		Normal Forms	Understanding what is Big Data.
Day3		Sequence Diagram	Provision ADLS Gen2
Day3		E-R Diagram	Creating containers,directory and uploading files in ADLS Gen2
Day4		Flowcharts	Provision Azure Databricks and its overview
Day4		Physical Database Design	Creating clusters in databricks.
Day5		Business Constraints – Primary key, Foreign Key, Candidate Key, Composite Key, Surrogate Key, Unique Key, Check, Default, Not Null	Widgets Utity
Day5		Index- Clustered and Non-Clustered	Widgets Utity
Day6			File System Utility,Understanding Cluster Configuration in Databricks
Day7	Revision		
Day8		Download and Install SSMS, SQL Server	Databricks CLI overview
Day8		SQL commands- DDL, DML, DCL, TCL	Service Principal
Day8		SQL commands- DDL, DML, DCL, TCL	Mount Point
Day9		SQL commands- DDL, DML, DCL, TCL	How to create Mount Point using Azure Key Vault
Day10		Select, INSERT, Update, Delete, Merge, Create, Alter, Truncate, Rename, Grant, Revoke, Commit, Rollback, Save point, Drop	Connect ADLS Gen2 from databricks using OAuth 2.0 with an Azure service principal
Day10		Query Execution sequence/Order	Connect ADLS Gen2 from databricks using Shared access signatures (SAS)
Day11		Clause, Operators, Predicates in SQL	Connect ADLS Gen2 from databricks using Account keys
Day11		LIKE, IN, LIMIT, TOP, CASE, INSTRING, SUBSTRING, AVERAGE, COUNT, MAX, MIN, SUM	Read and write data in Azure Databricks
Day12		UPPER, LOWER, EXIST, EXTRACT, Wild Card Operators Etc.	Data processing in Azure Databricks
Day12		SUBQUERIES - Nested Subqueries, Corelated Subqueries, Inline View	Working with DataFrames in Azure Databricks
Day13		JOINS – Self Join, Inner Join, Left Outer Join, Right Outer Join, Cross Join	Working with different types of file formats like CSV,Parquet,Avro,Orc
Day14	Revision		
Day15		SET Operators – Union, Union All, Minus, Intersect	
Day15		CTE (Common Table Expression)	More dbutils commands
Day15		Window Function – Rank, Dense Rank, Row number, Lead, Lag, First Value, Last_Value, Percent Rank	Databricks Runtime
Day16		Pivot and Unpivot	Concept of Delta Table
Day17		Rollup, Cube, Grouping Sets	Deep Dive into Delta Tables
Day18		Views, Materialized View	Managed and Unmanaged Table in Databricks
Day19			Table batch reads and writes
Day19			Table delete, update, and merge
Day20			Different types of cluster in databricks
Day20			Scheduling jobs in Databricks Notebook
Day21	Revision		
			Azure Key Vault
Day22	Basics of Python programming, working with Python interpreter identifiers		Provisioning of Azure Key Vault
Day22	keywords constants variables,		How to securely store Keys and Password in Azure Key-Vault
Day22	types of operators precedence of operators data types		How to access Keys and Password for Azure Key-Vault
Day23	mutable and immutable data types statements		How to connect AKV from ADB(Azure Databricks) and read the secrets of AKV
Day23	expressions		Secret Scope and databricks utility of secret scope
Day24	evaluation and comments input and output statements data type conversion		How to connect AKV from ADF(Azure Data Factory) and read the secrets of AKV.
			Azure Data Lake Gen2(ADLS Gen2)
Day25	selection (decision) and repetition (iteration)		Blob Storage
Day25	Selection: if, if-else, and nested if statement, indentation Repetition: for, while, and nested loops, break, continue		Azure DataLake Gen1
Day26	Chapter 3: Functions		Azure DataLake Gen2
Day26	Introduction to		Installation of Azure Storage Explorer and connection with ADLS Gen2
Day27			Difference between Blob Storage,ADLS Gen1 and ADLS Gen2
Day27			How do we store data in ADLS
Day27	Introduction to functions		How to read/write data in ADLS Gen2 from databricks
Day28	Revision		
Day29	need of functions User defined functions: passing arguments to a function, returning values from functions,		
Day29	scope of variables		
Day30	Standard library:, using built-in functions,importing modules-math, random, statistics, creating and importing user defined module		ETL CONCEPT
Day31	Periods Strings: initializing strings and accessing strings string operations		Data Warehouse Concept
Day32	built-in functions for string manipulation string traversal		Difference between Database,Data Warehouse and Data Lake
Day34	Periods Lists: list operations - creating, initializing, traversing and manipulating lists list methods and built-in functions		OLAP vs OLTP systems Concept
Day33	nested lists		Fact Table and Dimension Table
Day35	Revision		
Day36	list as argument to a function		Data Modelling-Star Schema vs Snowflake Schema Concept
Day37	Tuples: Creating, initializing, accessing elements, tuple assignment, operations on tuples, tuple methods and built-in functions, nested tuples.		
Day38	Dictionary: concept of key-value pair, mutability, creating, initializing, traversing, updating and deleting elements; dictionary methods and built-in functions.		
Day39	Exception Handling		SPARK
Day40	need of exception handling user-defined exceptions raising exceptions		Architecture of Spark
Day41	Try - except - else clause Try - finally clause		Job,Stages and Task in Spark
Day42	Revision		
Day43	File Handling: text file, file types open and close files		SparkSQL concepts,Pyspark concepts
Day44	reading and writing text files		Transformation and Action
			Lazy Evaluation
Day45,Day46-Day47	Modules		rdd vs dataframe vs dataset
			Immutability Concept
			operation on RDD
Day48	Packages		operation on dataframe
Day49	Revision		Difference file formats like parquet,avro,orc and their differences
			Understanding Shuffling in Spark
Day50-Day51-Day52	Regular Expressions		Narrow Transformation
			Wide Transformation
			Accumulator
			Broadcast variable
Day53-Day54-Day55	Multi Threading		Partition By
Day56	Revision+Start Thinking on Resume Preparation		
Day57-Day58	Connecting Databases with Python		partition pruning,repartition and coalesce
Day59			cache vs persist
			Azure SQL Server
Day60			Provisioning of Azure SQL Server
Day60			Installation of SSMS and connection with Azure SQL Server
Day61			Installation of Azure Data Studio and connection with Azure SQL Server
Day61			Connect Azure SQL server from databricks
Day61			Connect Azure SQL server from Azure Data Factory
			Azure Data Factory(ADF)
Day62			Pipelines
Day62			Activities
Day62			Datasets
Day62			Linked Services
Day63	Revision+Prepare first round of resume		
Day64			Integration Runtimes
Day64			AutoResolve Infrastructue Runtime
Day64			Self Hosted Infrastructure Runtime(SHIR)
Day64			SSIS Infrastructure Runtime
Day65			Pipeline Run
Day65			Scheduling and Trigger of pipeline
Day66			Parameters
Day66			Variables
Day66			Orchestration of Pipeline in ADF
Day67			Triggers
Day67			Schedule Trigger and its System Variables
Day68			Event Based Trigger and its System Variables
Day69			Tumbling Widow Trigger and its System Variables
Day70	Revision+Make changes in the resume if required		
Day71-Day72-Day73			Data Flow Pipeline
Day74			Pipeline Level Parameters
Day75			Global Parameters
Day76			Creating Dynamic Pipeline creation
Day77	Revision+Make changes in the resume if required		
Day78			How to monitor ADF data pipelines?
Day78			ARM Template
Day79			Copy one table from My SQL(on-premise) to Azure SQL Server.
Day79			Copy filtered data of one table from My SQL(on-premise) to Azure SQL Server.
Day80			Copy All tables from My SQL Server(on-premise) to Azure SQL Server
Day80			Copy selected tables from My SQL Server(on-premise) to Azure SQL Server
Day81			Pipeline for Incremental Load
Day82			Copy all tables from Azure SQL Database to ADLS Gen2
Day83			Copy selected tables from Azure SQL Database to ADLS Gen2
Day84	Revision+Make changes in the resume if required		
Day85			Error Logging of ADF pipelines in Azure SQL Server
Day86			Azure Logic App
Day86			How to send mail alert for pipeline failure and success using logic app.
Day87			Deployment
Day87			Introduction to Deloyment
Day87			DEV,UAT and PROD envrionment
Day88-Day89			Azure DevOps
Day90			Using CI/CD pipelines for deployment of Azure Databricks and Azure Data Factory
Day91	Revision+Make changes in the resume if required		
Day92-Day93	Medallian Architecture		
Day94-Day95-Day96-Day97	Creating data pipelien using all major components		
Day98	Revision+Make changes in the resume if required		
Day99	Final Resume preparation/Mock Interew		
Day100	Final Resume Preparation/Mock Interew		

GET SET READY TO GIVE INTERVIEW!! ALL THE BEST!!

Tips:-
1)Resume preparation is a continuous process.It will take multiple iteration of resume bulding to build a robust resume.
2)Consider each failed interview as the mock interview and prepare the set of question you were not able to answer before sitting in the next interview.
3)Keep on practicing different sets of question in Python and SQL during the entire process of learning from any portal.
4)Continuous practice and revision is the key to learn technology and crack interviews.
5)That day will never come e