	All Rounder Preparation									
This is a basi	c data Science roadmap with links and instructions as to l Use it to	how to become a proper data scientist. It may inlu byour own personal discretion and keep it updates	de neccessary and sometimes not so neccessary but helpful topics to work upon. d when neccessary.	Dat	te Comple	eted	Perso	nal Difficu	ilty 1-5	Notes
Serial Number	Topic Name	Platform I	Ratform 2	Hemant Gamin	Prateek Dilaware	Tanmay Somani	Hernant Gamin	Prateek Dilaware	Tanmay Somani	ni Hemant Gamin Proteek Dilaware Tanmay Somani
11 12	Matrices and Linear Algebra Hash Function, Binary Tree ,O(n)	Linear Algebra and Matrices Hash Functions , Binary Tree	www.geeksforgeeks.org/introduction-to-hashing-data-structure-and-algorithm-to-	storiols/	09-02-2024	22-01-24 23-01-24		3	3	Storted, to be continued Binary search tree remains Linkedin lists is a prerequisite, binary tree again
13 14	Relational Algebra , DB Basics Inter, Outer, Cross, Theta Join	Belational Algebra GATE Notes, Db for Beginners Joins				25-01-24 25-01-24			2	very basic though must remember the notations easy enough, just remember cross join
15 16	Tabular Data	Bosics of CAP				26-01-24 26-01-24			1	ACID is different.
17 18 19	Entropy Data Frames and Series Sharding	Entropy DateFrame and Series Ob sharding				29-01-24 30-01-24			2	ofso see replication
110	OLAP	OLAP Multidimensional Data Model				31-01-24 31-01-24			1	diso read OLTP
112 113	ETL	Reporting Ve BI , Bi Ve Ba				01-02-24 01-02-24			1 2	easy definition and fundamental differences
114 115	NOSql	JSON VS XML, Json Vs Xml NoSOL				01-02-24 01-02-24			2	need deeper understanding
116 117	Vendor Landscape	Regex IDE_VCS_Gradle, Maven, or Webpack_Libraries,f							2	connot understand at all
118 2 21	Statistics	IDE_VCS_Gradia, Moven, or Webpack_Libranes,	-			01-02-24			2	Personal Choices Need to start working on datasets for now onto sit
	Descriptive Staticatics (mean,median,mode,range, SD)		EDA.			02-02-24			1	clearly understood do it again.
2.4 2.5										
26 27	Bayes Theorem	Probability Theory Bayes Theorem				04-02-24			2	
28 29 210	Random Variables Cumulative Distribution Function Continuous Distributions									
2 II 2 II 2 I2	Skewness ANOVA									
213 214	Probability Density Function Central Limit Theorem									
2 15 2 16	Monte Carlo Method Hypothesis Testing									
217	P value Chi ³ Test									
2 19 2 20 2 21	Estimation Confidence Interval MLE									
2 22 2 23	MLE Kernel Density Estimate Regression									
2 24 2 25	Covariance Correlation									
2 28 2 27	Pearson's Coefficient Causation									
2 28 2 29 3	Least-fit Euckedean Distance Programming									
31	Python Basics Working in Excel									
3 3 3 4	R Studio Setup Expressions									
35 36	IBM SPSS Variables									
37 38	Ropid Miner Vectors									
3 9 3 10 3 11	Matrices Arrays Factors									
3 12 3 13	Lists Data Frames									
314 315	Reading CSV/Raw Data Subsetting Data									
3 16 3 17		https://www.youtube.com/watch?v=aXlwWbU865	22							
318 319 320	Functions Factor Analysis									
3 21	Install Packages Exploratory Data Analysis Machine Learning									
41 42	What is Machine Learning? Numerical Variance									
43 44	Categorical Variance Supervised Learning									
45 46 47	Unsuperized Learning Concepts, inputs and Attributes Training, Input and Attributes									
48	Training & Test Data Classifier									
4 10 4 11	Prediction Lift									
4 12 4 13	Overlitting Bias and Variance									
414 415 416	Trees and Classification Classification Rate Decision Trees									
4 17 4 18	Boosting Naive Bayes Classifiers									
4 19 4 20	K-N-N Logistic Regression									
4 21 4 22	Ranking Linear Regression									
4 23 4 24	Perceptron Clustering Neural Networks									
4 25 4 26 4 27	Neural Networks Sentimental Analysis Callaborative Filetring									
4 28 5	Tagging Text Mining / NLP									
51 52	Corpus Entity Recognition									
53 54	Test Analysis UMA Term Document Motrix									
55 56 57	Term Document Matrix Term Frequency And Weight Support Vector Machines									
58 59	Association Rules Market Based Analysis									
510 511	Feature Extraction Using Mahout									
5 12 5 13	Using Welka Using NLTK									
514 515 516	Classify Text Vocabulary Mapping Tagging									
61	Visualisation Data Exploration in R									
62 63	Uni,Bi and Multi Variate Viz ggplotz									
64	Histogram and Re Tree and Tree Map									
66 67 68	Scatter PLot Line Charts Spatial Charts									
68 69 610	Survey Plot Timeline									
6 II 6 I2	Decision Trees D3 js	Dija								
6 13 6 14	InfoVis IBM ManyEyes									
6 15 7 71	Tableu Big Data MapiReduce Fundamentals									
71 72 73	Hadoop Components									
7.4 7.5	Data Replication principles Setup Hadoop									

76	Name and Data Nodes					
77	Job and Task Tracker					
78	MR Programming					
79	Sqoop : Loading Data in HDFS					
710	Flume Scribe for unstructed data				i	
7 m	9QI with pig				ĺ	
712	DWH with Hive				i	
713	Scribe chukwa for Weblog				1	
714	Using Mahout					
715	Zookeeper Avro					
716	Storm Hadoop					
717	Rhaddop Rhipe					
718	rmr					
719	Cassandra					
720	MongoDB,neo4j					
8	Data Ingestion					
81	Summary of Data Formats					
82	Data Discovery				i	
83	Data Sources and Acquisition				1	
84	Data Integration					
85	Data Fusion				-	
86						
	Transformation and Enrichment					
87	Principal Component Analysis					
88	Data Survey					
89	Google OpenRefine					
810	how much Data?					
811	Using ETL				i	
9	Data Munging					
91	Dimensionalities and Normalities Reduction					
92	Normalization				 i	
93	Data Scrubbing					
0.4						
94	Handling Missing Values					
95	Handling Missing Values Unbiased Estimators					
95 96	Handling Missing Values Unbiased Estimators Bining Sparse Values					
95 96 97	Handling Missing Values Unbiased Estimators Bining Sparse Values Feature Extraction					
95 96 97 98	Handling Missing Values Unblassed Estimators Bining Sparse Values Feature Estraction Denoising					
95 96 97 98 99	Handling Missing Values Unbiosed Estimators Bining Sparse Values Feature Estraction Denoising Sampling					
95 96 97 98	Handling Missing Values Unblassed Estimators Bining Sparse Values Feature Estraction Denoising					
95 96 97 98 99	Handling Missing Values Unbiosed Estimators Bining Sparse Values Feature Estraction Denoising Sampling					
95 96 97 98 99	Handling Missing Values Unbiased Estimators Bining Sparse Values Feature Estraction Denoising Sampling Stratified Sampling					
95 96 97 98 99	Handling Missing Values Unbiased Estimators Bining Sparse Values Feature Estraction Denoising Sampling Stratified Sampling					
95 96 97 98 99 910	Honding Making Values Unbiased Estimators Bining Sparse Values Feature Energian Denoising Sampling Sampling Sampling Sampling Sampling Sampling Sampling					
95 96 97 98 99 910 911 101	Hendling Making Visibuse Verbissed Stamotors Bening Sparse Visibuse Feature Education Develoing Scorping Storolist Sampling Storolist Sampling Principal Component Analysis MS Seed of Costo Analysis MS Seed of Costo Analysis					
95 96 97 98 99 910 911	Handling Making Values Unclosed Estruction Biring Sparse Values Feature Enroction Denoising Sampling Stantified Sampling Principal Component Analysis MS Sized w/ Data Analysis India					
95 96 97 98 99 910 911 10 101 102 103 104	Preciding Making Working Unbiased Stamptons Bring Sporse Viviaes Fedura Entraction Devicing Scorping Stampling Stampling Stampling Stampling Stampling Michael Stampling Micha					
95 96 97 98 99 90 910 911 10 101 102 103 104	Heradini gifasing vigorase Unbiased Esternature Bining Sporae Viniase Feature Entraction Devicining Sompling Stratification Sompling Stratification Phinosocion Component Annalysis Instruction As Execution Analysis of Southern Javan Jaymon R. Bascolomine Weaturolimine Sporad More Moscope					
95 96 97 98 99 910 911 10 101 102 103 104 105	Preciding Making Working Unbined Stimmtons Bining Sporce Visions Finding Sporce Visions Federace Entraction Devoking Sompley Strottled Sampleying Strottled Sampleying Whitego Control Control Assistance of Cotta Anapagina Strottle Making Assistance of Cotta Anapagina Strottle Assistance of Cotta Anapagina Strottle Making Assistance of Cotta Anapagina Strottle Assistance o					
95 96 97 98 99 910 101 102 103 104 105 106 107	Heradini platesing visuales Unbiased Esternatures Bining Sporae Visuales Feature Endoction Devicining Sompling Stratification Stratification Devicining Sompling Stratification Sompling Phinogra Component Annahysis Institute Sompling Mission Stratification Stratification Association Stratification Association Stratification Association Stratification Visual Stratification Hostoop Sporik Storm					
9 5 9 6 9 6 9 7 9 8 9 9 9 10 9 11 10 2 10 3 10 4 10 5 10 6 10 7 10 8	Preciding Melang Visions Unbiased Stamptors Bining Sporse Visions Bining Sporse Visions Feature Entraction Develoing Sampling Stratified Sampling Stratified Sampling Philospic Component Analysis Mill Social will Dato Analysis Toolina John Share Sampling Mill Social Will Dato Analysis Toolina John Share Sampling Mill Social Will Dato Analysis Toolina John Share Sampling Soon Storm Financial Calculation					
9 5 9 6 9 7 9 8 9 9 9 10 9 11 10 1 10 2 10 3 10 4 10 5 10 6 10 7 10 8	Heradini planning vintures Unbiased Esternatures Bining Sporae Vintues Feature Endoction Devicining Sompling Straffield Sompling Straffield Sompling Phinogol Component Annahysis Instituted Sompling Mission of Component Annahysis Institute Sompling Mission of Component Annahysis Institute Sompling Mission of Component Annahysis Johnship Mission Mission Sompling Sporit Sporit Straffield Sompling Flums-Schalchakon Mistri Telenation Googeneit Mission Straffield Sompling Mission					
9 5 9 6 9 7 9 8 9 9 9 10 9 11 10 1 10 2 10 3 10 4 10 5 10 6 10 7 10 8 10 9 10 10 10	Preciding Melang Visions Unbiased Stamotors Bining Sporse Visions Feature Entraction Developing Sampling Stratification Stamotors Developing Stratification Stratification Developing Stratification Stratification Developing Stratification Stratification Developing Stratification Stratification Developing Meland Stratification Developing Stratification Developing Stratification Developing Stratification Stratification Stratification Stratification Stratification Stratification Developing Stratification Stratification Developing Stratification					
95 96 97 98 99 910 911 101 102 103 104 105 106 107 108 109 1010	Preciding Making Working Unbload Stamptons Bring Sporse Violes Fedora Entorction Devicing Stamping Making					
95 98 98 99 910 911 101 102 105 105 105 105 105 105 105 105 105 105	Preciding Melang Visions Unbiased Stamotors Bining Sporse Visions Feature Entraction Developing Sampling Stratification Stamotors Developing Stratification Stratification Developing Stratification Stratification Developing Stratification Stratification Developing Stratification Stratification Developing Meland Stratification Developing Stratification Developing Stratification Developing Stratification Stratification Stratification Stratification Stratification Stratification Developing Stratification Stratification Developing Stratification					
95 96 97 98 99 910 911 101 102 103 104 105 106 107 108 109 1010	Preciding Making Working Unbload Stamptons Bring Sporse Violes Fedora Entorction Devicing Stamping Making					
95 98 98 99 910 911 101 102 105 105 105 105 105 105 105 105 105 105	Heradining Makang Walanas Unbiased Sistematura Bering Spores Virules Fedura Potoccion Devicining Soveries Soveries Soveries Marcal Control Control Soveries Marcal Control Marcal Control Marcal Control Marcal Control R. Education Stories Marcal Control R. Education Stories Marcal Control Soveries Marcal Control Soveries Marcal Control Soveries Marcal Control Soveries Marcal Control Marc					
95 98 97 99 99 99 99 99 99 99 99 99 99 99 99	Preciding Messing Visions Unbiased Estimators Bining Sporse Visions Februs Entraction Devicing Sompleig Strottled Sompleig Strottled Sompleig Strottled Sompleig Strottled Sompleig Mining Component Analysis Mining Control Analysis Mining Control Analysis Mining Control Analysis West Control Mining Control					
95 96 97 98 99 90 90 90 90 90 90 90 90 90 90 90 90	Heradining Makang Walanas Unbiased Sistematura Bering Spores Virules Fedura Potoccion Devicining Soveries Soveries Soveries Marcal Control Control Soveries Marcal Control Marcal Control Marcal Control Marcal Control R. Education Stories Marcal Control R. Education Stories Marcal Control Soveries Marcal Control Soveries Marcal Control Soveries Marcal Control Soveries Marcal Control Marc					

S . No	Name	Link	Note
1	Learn Dancing	https://www.youtube.com/playlist?list=PLGN5InbZpwnnLWEHfMMqdKOBimH1aXOZp	
2	Basic Data Science Pathway	https://t.me/machinelearning_course	
3	Aptitude Test Practise	https://www.javatpoint.com/aptitude/quantitative	
4	Machine learing guide		
5	Leetcode DSA	<u>Leetcode Blind 75</u>	
6	Foage Virtual Internships	https://www.theforage.com/	
7			
8			
9			
<u>S.No</u> .	Project Name	Links	Note
1	Al image classification		
2	Intelligent Traffic Management System and Challan		
3	Attendance Image classification		
4			
5			
6			
7			
8			
9			
10			

3 newsletters subscribed, Learn GitHub		Concepts of SQL	Documentati on 2 [85%], 2 Presentations, Story using 50 new learnt words, 1 research ideas	Concepts of Python (Basic, Data Structures, Libraries, Pandas, numpy, matplotlib, seaborn, tkinter, pytorch or scikitlearn) [90%]; pip, Project 3; WPM expected to be near 35, 60 percent score using careerflow	Resume Creation		Concept of Exc	All resources Completed!!!,		
3 Project Ideas, Learn GitHub	1 Project, 2 newsletter subscribed ,	Concepts of SQL	1 Project, Story using 50 new learnt words Story using	1 Project, 1 SQL Project, WPM expected to be near 35, Concepts of Python(Basic, Data Sctructures). 60 percent using careerflow	Resume Creation	5 Speeches , 2 :3 ; 5:2	Concept of Excel	All resources Completed!!!, 1 Internship		
Linkedin Post, 7 Instagram posts, 1 Minor projects	Linkedin Post, 7 Instagram posts	Concepts of SQL,Linkedin Post,7 Instagram posts	50 new learnt words, Linkedin Post, 7 instagram posts , 1 Minor project	WPM expected to be near 50, Linkedin Post, 7 instagram post	Resume Creation, Instagram post	Linkedin Post, 7 Instagram post	Linkedin Post, 7 Instagram post, 1 Minor Project	All resources Completed!!!, 1 Internship, 1 Major Project	7 Instagram post	7 Instagram post

S No. Date Name of the company (Company waters Company Linkedin Source First mat/ first message Follow up 1 Follow up 2 Conclusion										S No. Date Name of the conf Company websit Company Linked Source First mail fast in Police up 1 Police up 2 Conclusion									INTERNAME										
S No.	Date	Name of the company	Company website	Company Linkedin	Source		Follow up 1	Follow up 2	Condusion	S No.	Date	Name of the con	Company websit	Company Links	Source	First mail/ first m	Follow up 1	Follow up 2	Conclusion	S No.		Name of the com		Company Linked	Source	First mal/ first m	Follow up 1	Follow up 2	Conclusion
																												_	
										_																		_	_
					_		_		_	_			_					_	_	_									_
					_					_			_					_							_			_	_
	_			_	_		_		_	_	_		_					_	_	_				_	_		_	_	_
				_	_					_	_		_					_	_						_			_	_
	_			_	_				_	_			_					_	_	_					_			_	
					_					_	_														_			_	_
										_																		_	
																												_	