


# SMART-ED

- Education Made Smarter -

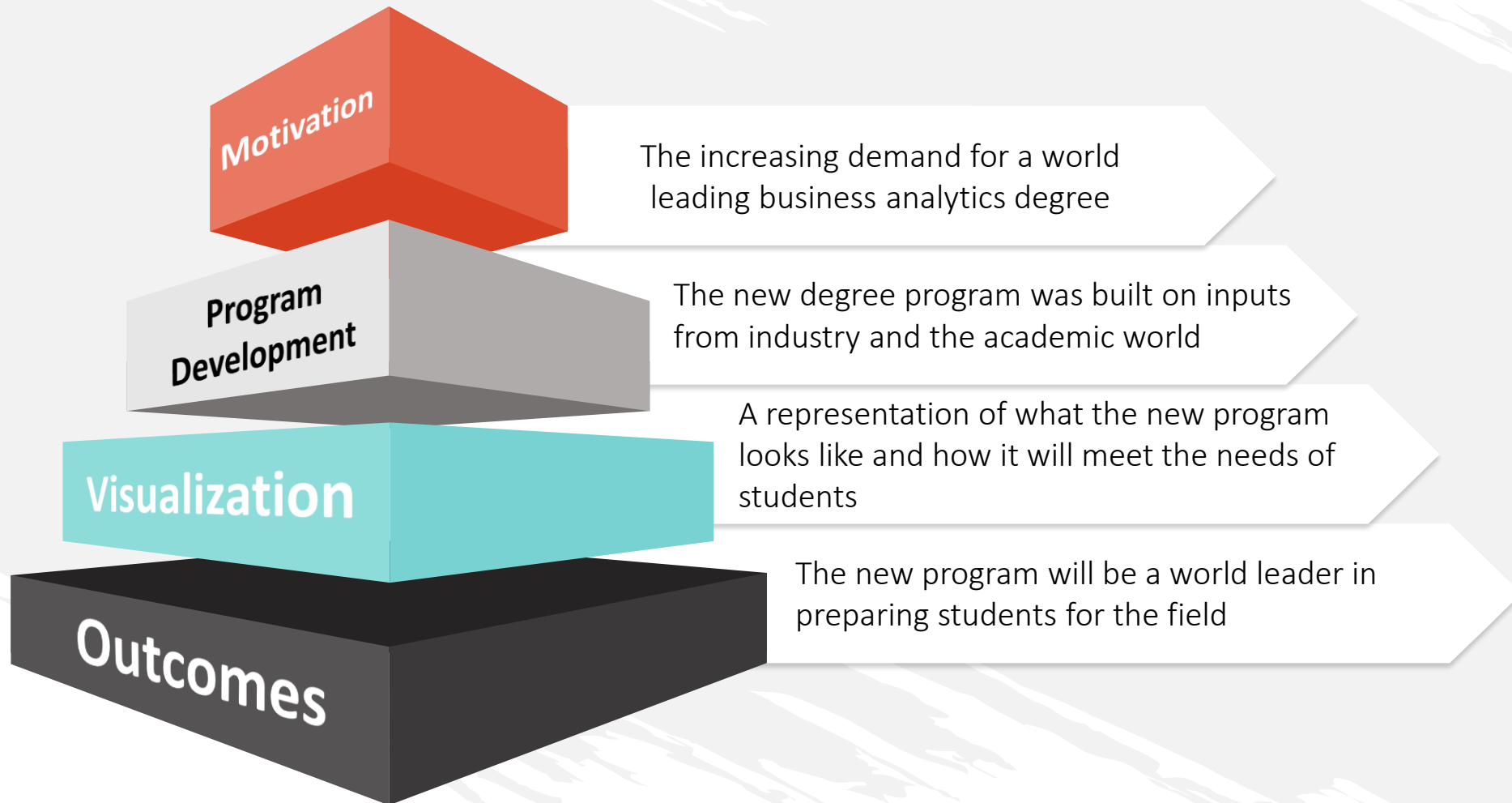


MIE 1624 Winter 2020 Group 11 – Startup Project

- Simon Faux | Akshat Mathur | Rami Al-Sahar | Gautam Dawar



# Business Ideology



# A Program For Students, Designed By Industry



## In Demand Skills

Data extracted from over 300 data scientist job postings



## Leading Courses

We picked top Universities which designed business focused courses



## Data Analysis

Used Novel models and scoring strategies to create best course for you!



## Our Program

A curriculum based to prepare students to become industry leader in Analytics

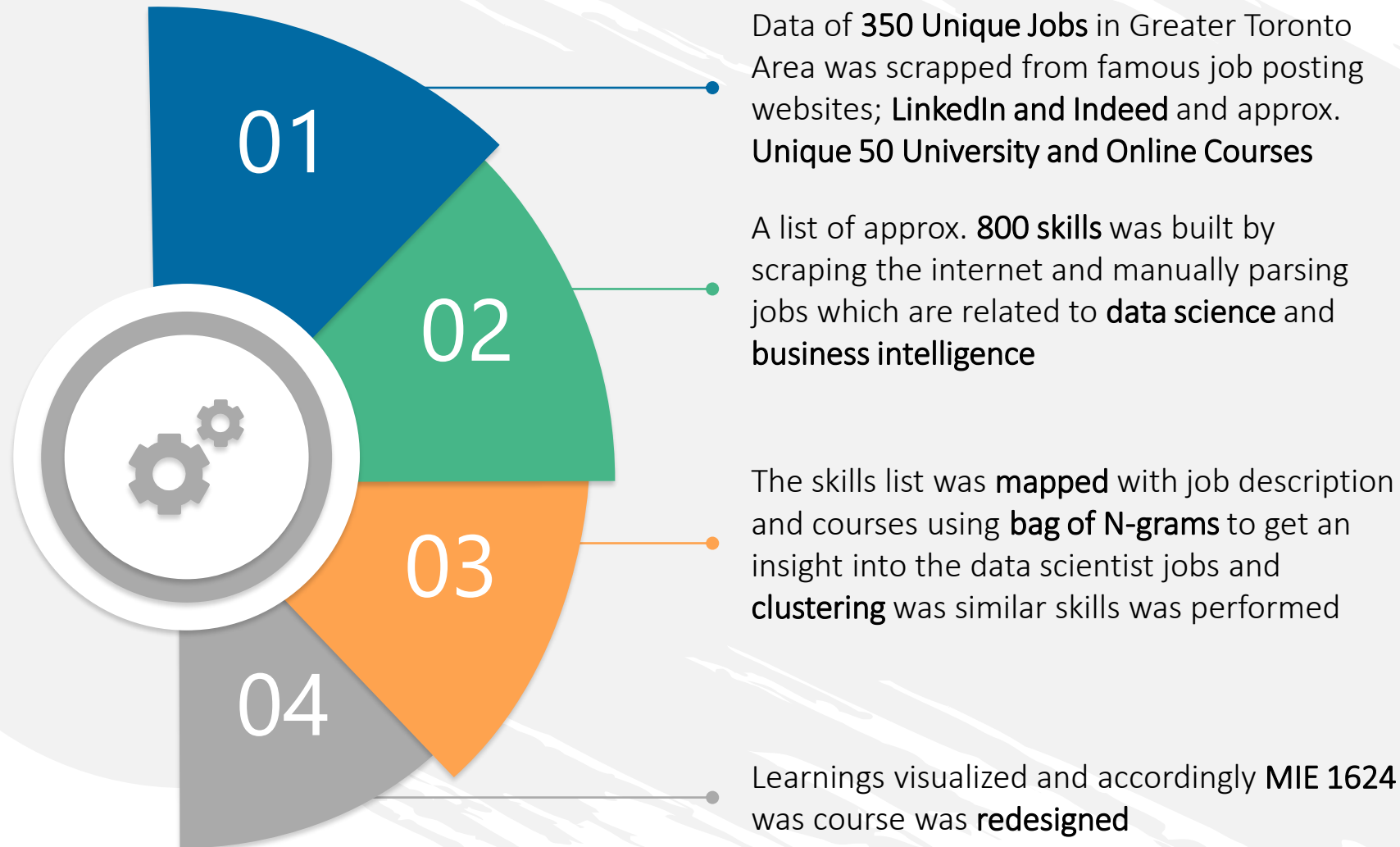


# Part 1:

# Job Skills Analysis and Redesign MIE 1624

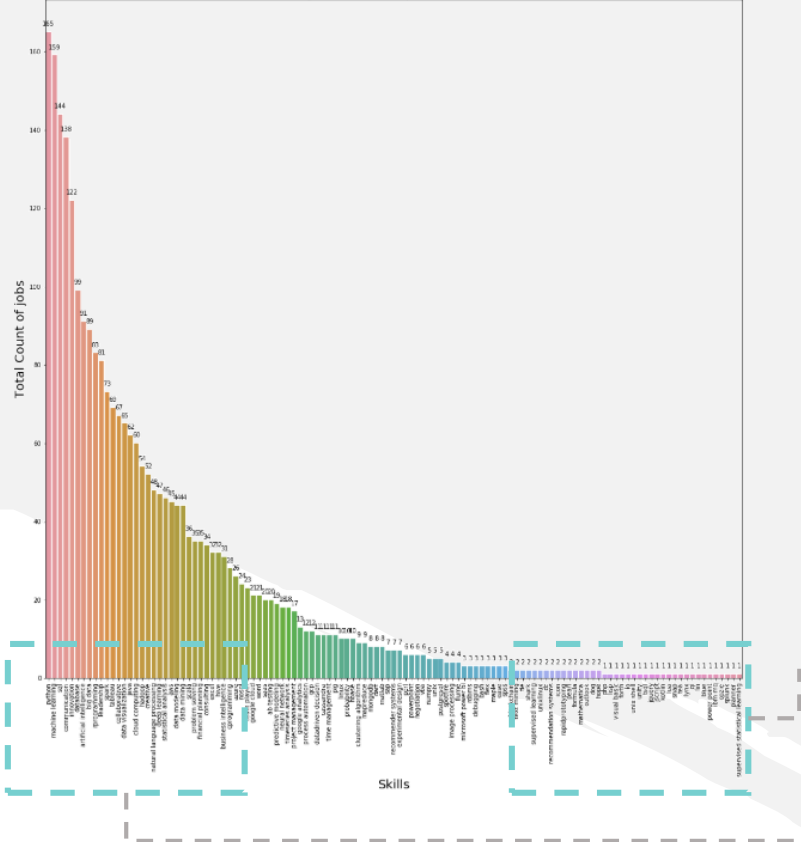


# Part 1: Model Functioning Overview

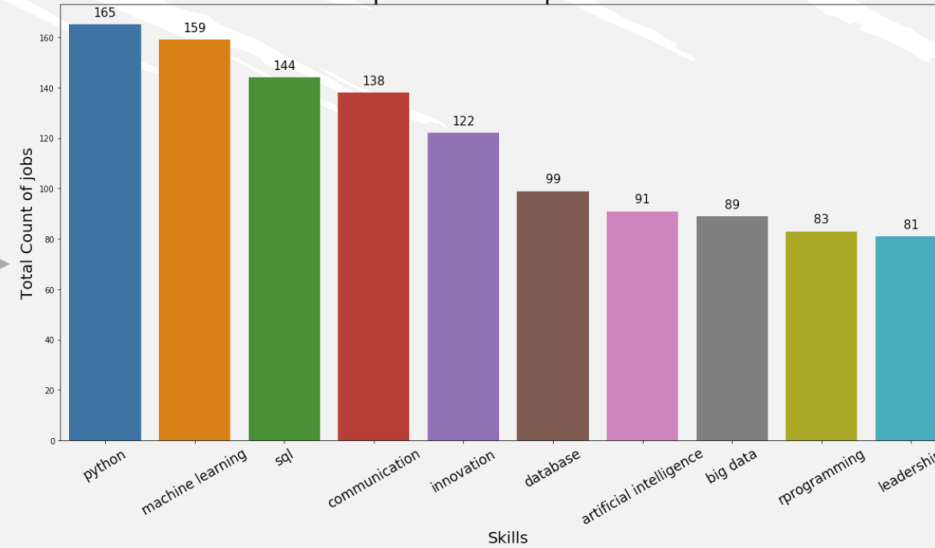


# Part 1: Skill Insights

Job Count per skill



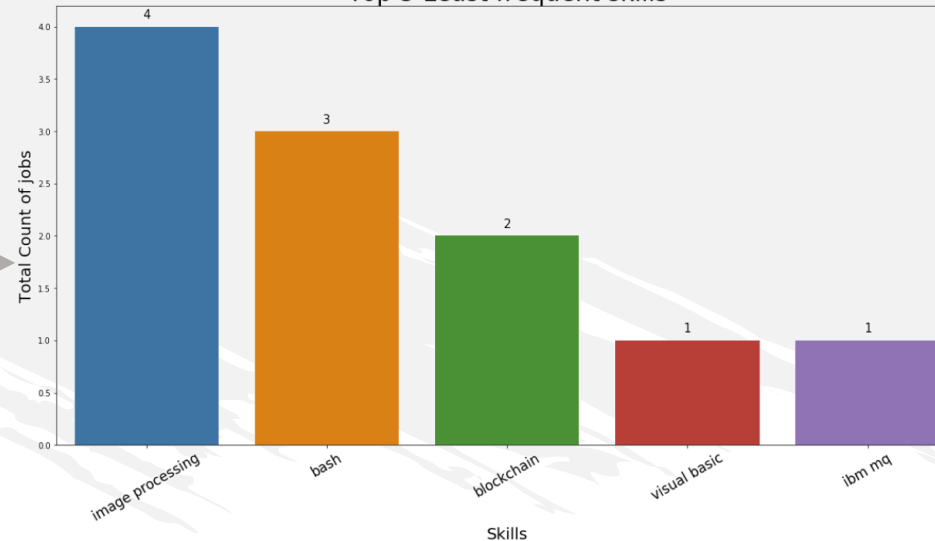
Top 10 most frequent skills



## Top 10 Most Frequent Skills

Stats and coding skills like **python**, **r**, **Artificial Intelligence** and **DBMS skills** like **sql**, **database**, **big data** and **soft skills** like **communication** and **leadership** are **most popular** among employers for Data Scientists

Top 5 Least frequent skills



## Top 5 Least Frequent Skills

**Image Processing**, **blockchain** and **Bash** are among least popular skills in demand by the employers for Data Scientist



# Part 1 Results: Fully Redesigned MIE 1624



## Week

## Outline

## Skills Gained

## Homework & Skills

1

Introduction to data science and analytics

Quantitative modeling | basic statistics | linear algebra | Python



#1 Assignment Python Basics | data exploration | Statistics | linear algebra

2

Modeling techniques, regression

Mathematical modeling | Linear regression | Logistic regression



#2 Assignment logistic regression | Tableau Dashboard

3

Data visualization and visual analytics

IBM Watson Analytics | Storytelling | analytical decision | Tableau



#3 Group Project NLP | collaboration | leadership | innovation | communication

4

Data mining & machine learning

Classification | Clustering | supervised machine learning | Data mining



#4 Assignment data mining | Hive | SQL | Spark

5

Optimization

non-linear optimization | constrained optimization

6

Simulation modeling

Monte Carlo | Random number generation

7

Cognitive computing & artificial intelligence

neural networks | natural language processing | Reinforcement learning

8

Enterprise Hadoop & Hadoop basics

HDFS | YARN | data management | big data | business case

9

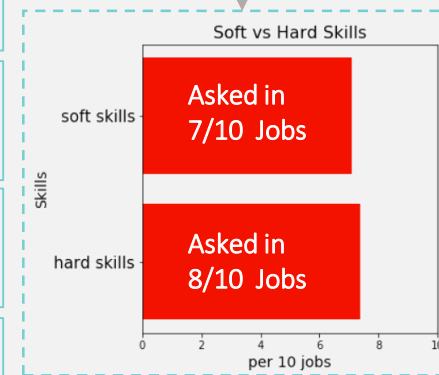
SQL: Hive and Impala

scripting | Spark | Hive | Impala | SQL: most important commands

10/11

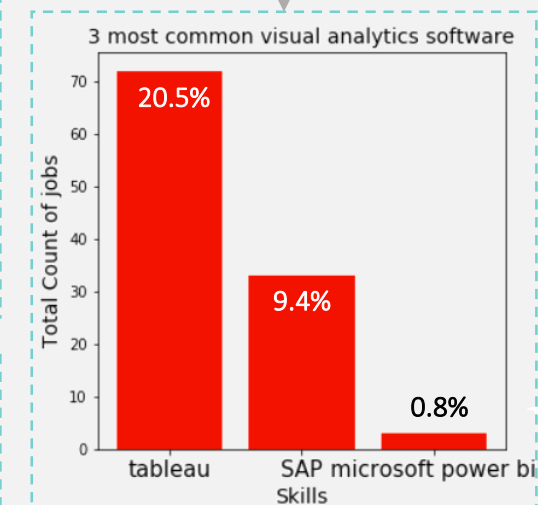
Spark, introduction

PySpark | time-series - Spark | Model development



43% approx.

Jobs require knowledge on SQL and Big - data



Existing Curriculum  
Modified Curriculum



# Part 1: Course Comparison

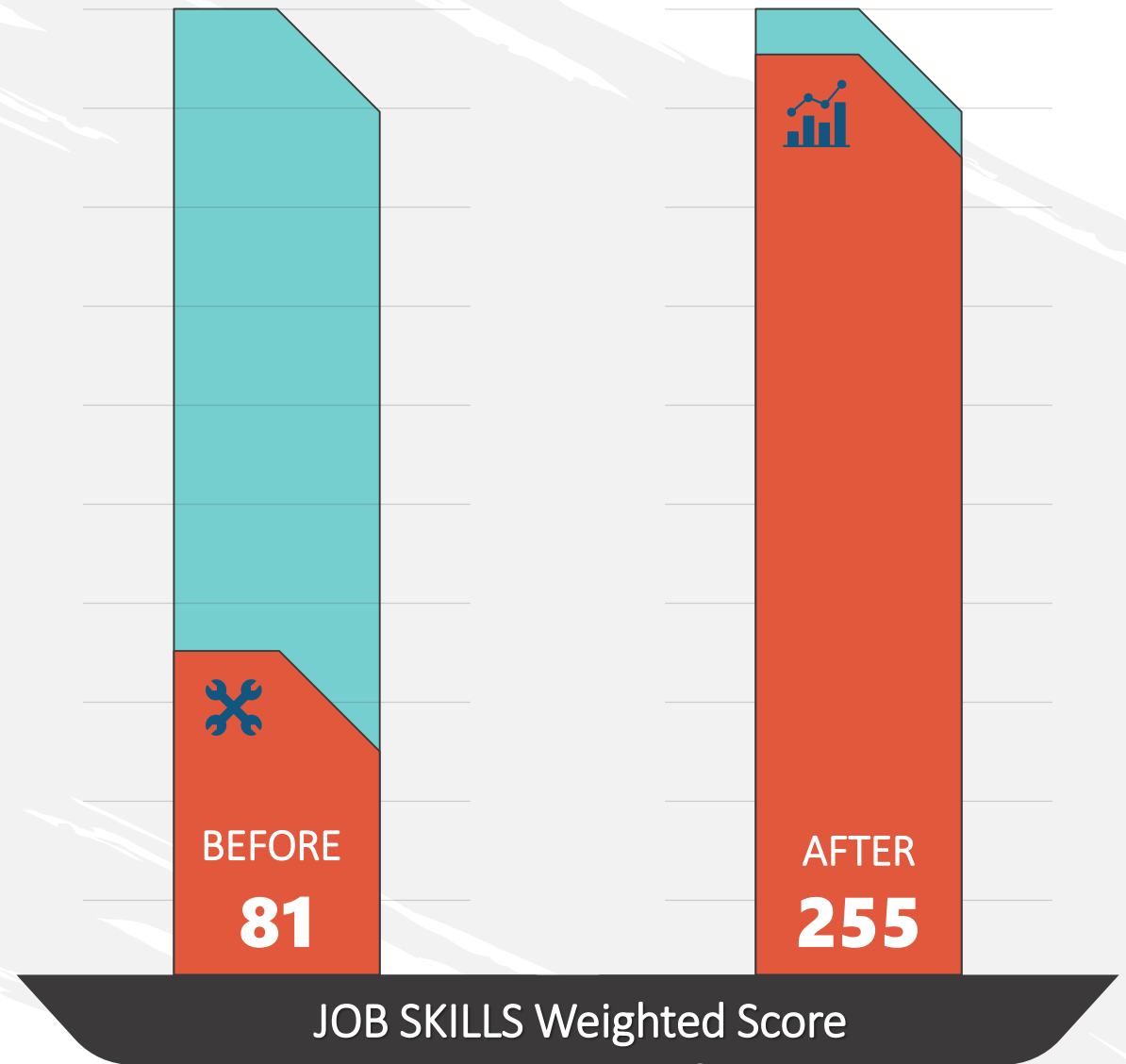
The following is a comparison between the number of jobs qualified for by studying the original MIE 1624 course vs. the updated MIE 1624 course.



# of jobs qualified for by taking the original course.



# of jobs qualified for by taking the redesigned course



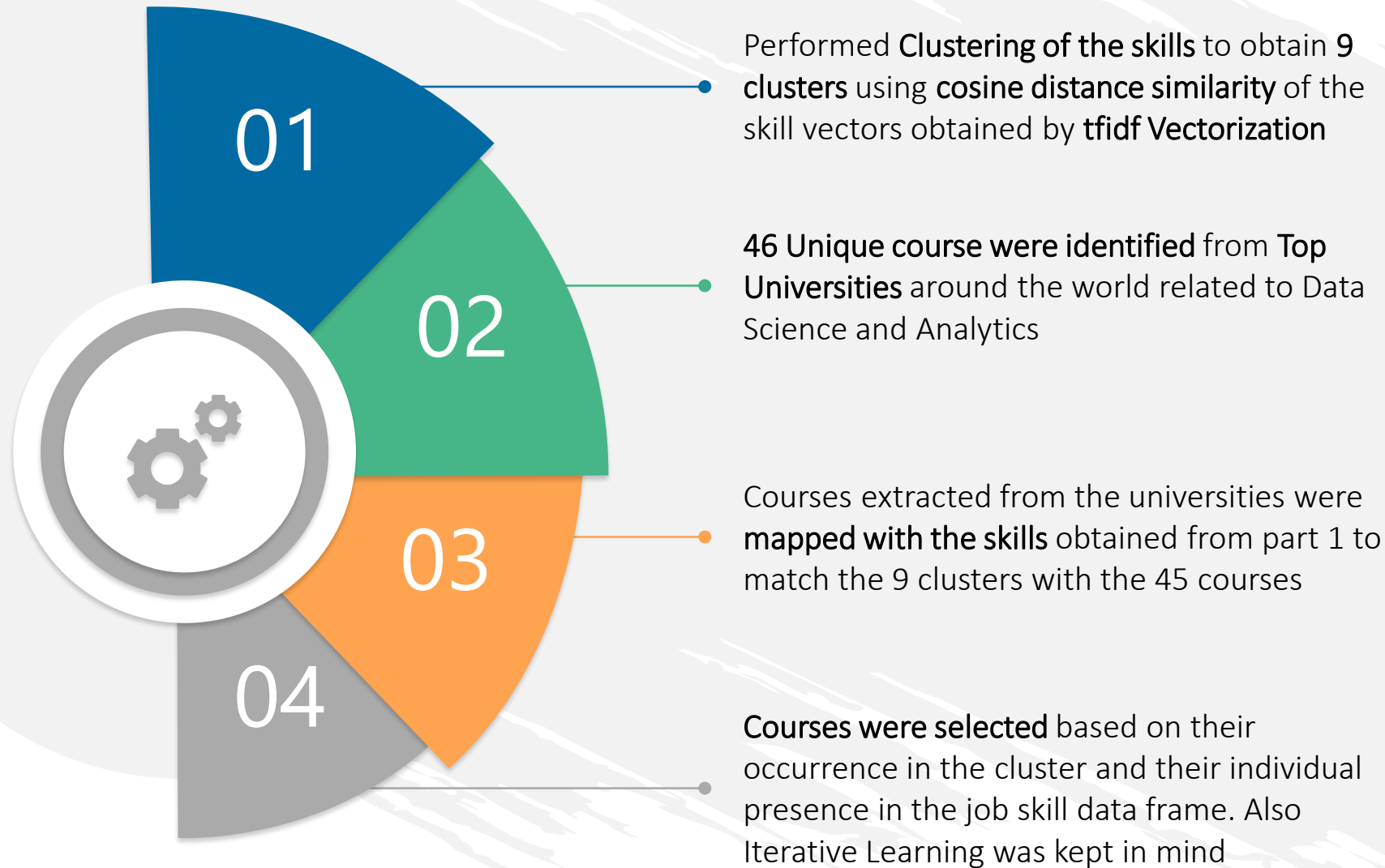
# Part 2:

## Masters Program Design –

### *Master of Data Science and Artificial Intelligence*



## Part 2: Model Functioning Overview



# Part 2: Skill Cluster Identification

01

'python', 'sql',  
'communication',  
'database',  
'rprogramming',  
'tableau', 'data  
visualization'

02

'saas', 'ab  
testing',  
'google  
analytics',  
'datadriven  
decision'

03

'nosql',  
'cassandra'

04

'big data',  
'spark', 'java',  
'hadoop',  
'scala', 'hive',  
'hbase'

05

'problem solving',  
'consulting',  
'business  
intelligence', 'word',  
'timeseries analysis',  
'project management',  
'process automation',  
'time management'

06

'innovation',  
'leadership',  
'collaborative',  
'creative', 'data  
modeling',  
'financial  
planning', 'excel',  
'team player'

07

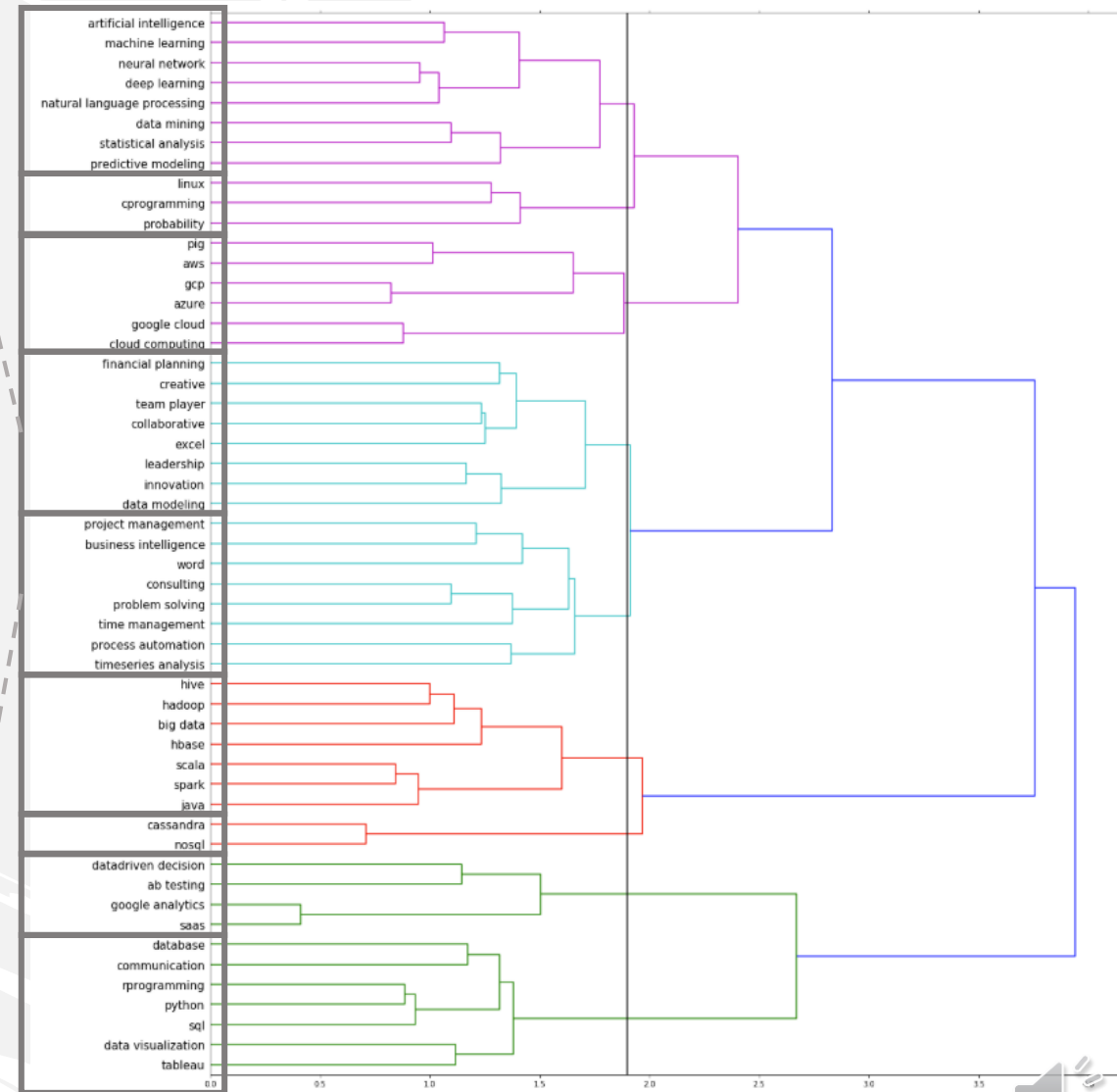
'cloud computing',  
'aws', 'azure',  
'google ', 'gcp',  
'picloudg'

08

'cprogramming',  
'probability',  
'linux'

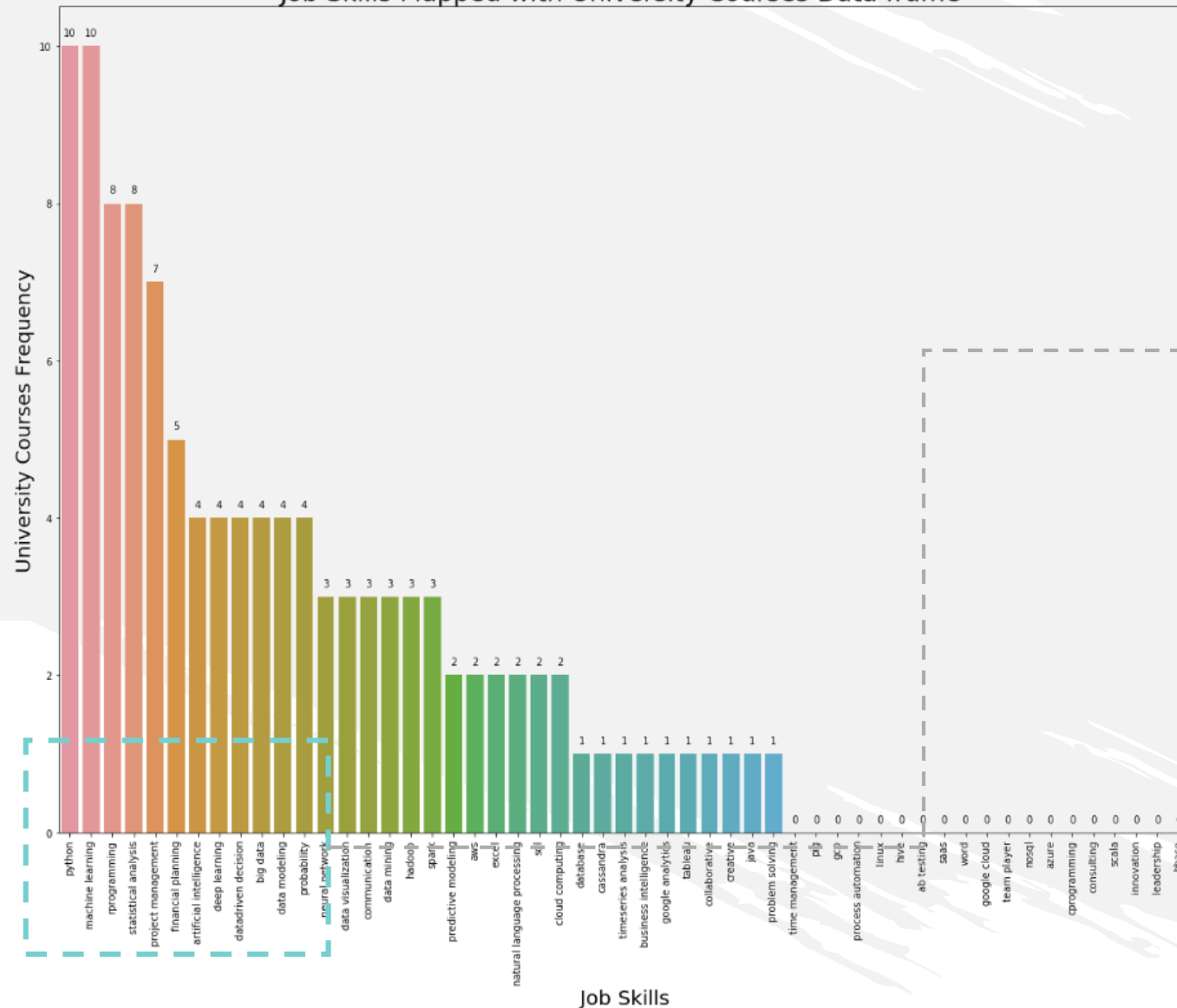
09

'machine learning',  
'artificial intelligence',  
'natural language  
processing', 'deep  
learning', 'statistical  
analysis', 'data mining',  
'predictive modeling',  
'neural network'

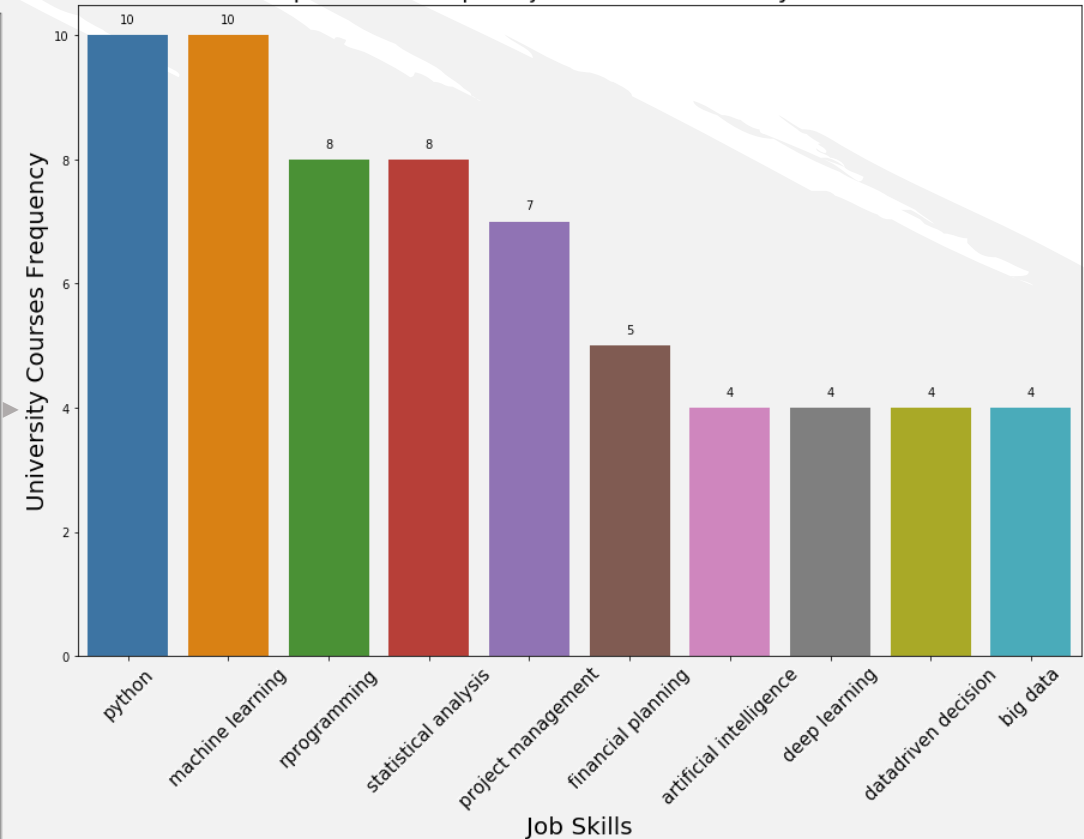


# Part 2: Job Skill Mapping with University Courses

Job Skills Mapped with University Courses Data frame



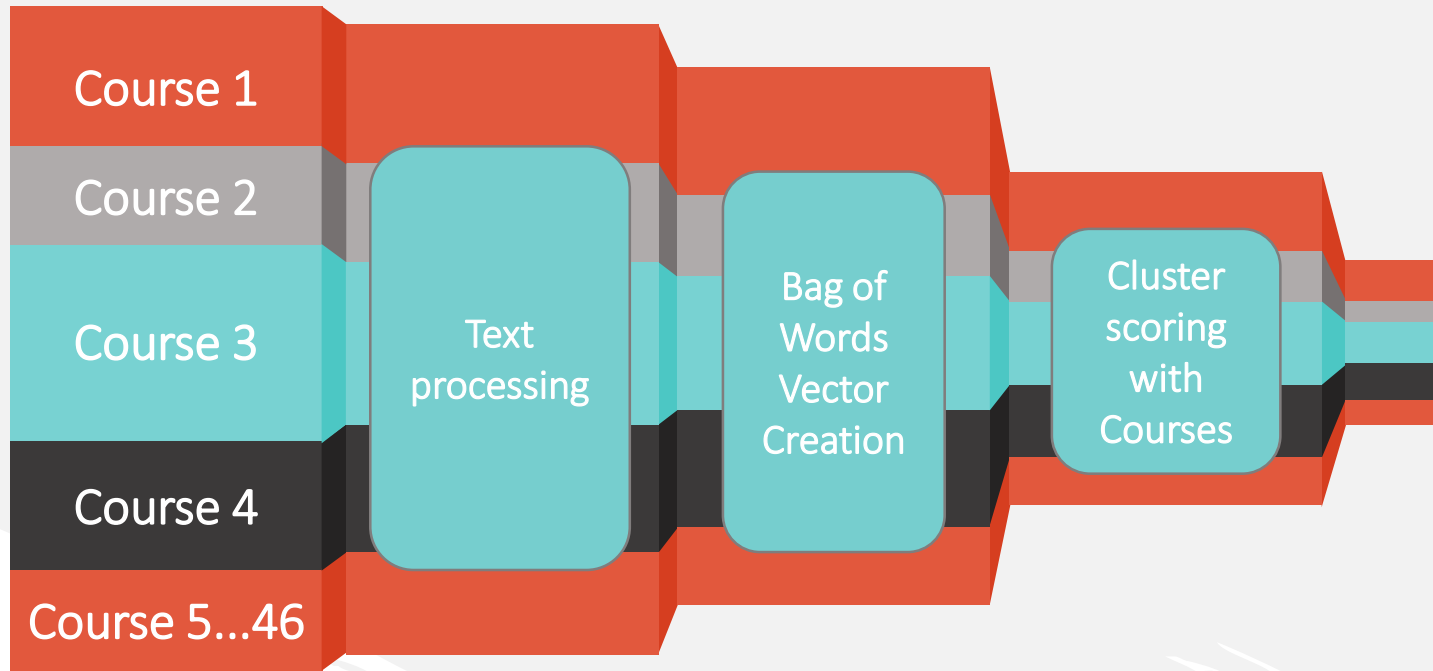
Top 10 Most frequent job skills in University Courses



## Top 10 Most Frequent Job Skills in University Courses

Skills like **python** | **Machine Learning** | **R** | **Statistical Analysis** | **Project Management** | **Financial Planning** | **Artificial Intelligence** | **Deep Learning** | **Data Driven Decision** | **Big Data** are quite common

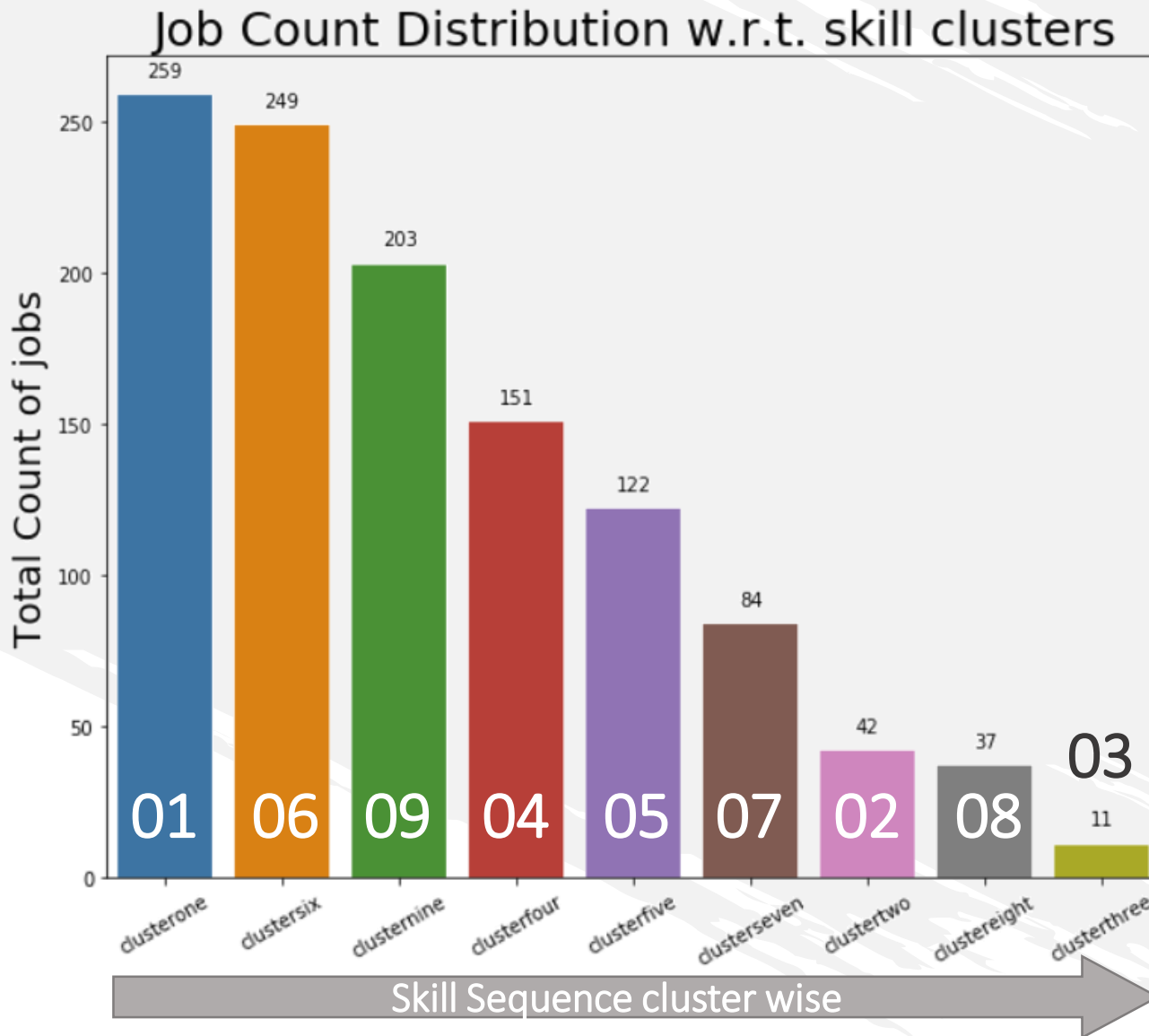
## Part 2: University Courses Mapping to Job Skill Clusters



cluster_number		Title
11	1	Data Visualization
1	2	Research Design and Application for Data and Analysis
7	3	Deep Learning in the Cloud and at the Edge
9	4	Machine Learning at Scale
8	5	Statistical Methods
2	6	Statistics for Data Science
3	7	Fundamentals of Data Engineering
4	8	Applied Machine Learning
21	9	Leveraging AI and Deep Learning Tools in Marketing (RSM8521)



## Part 2: Skill Cluster Sequencing based on Job Skills Distribution



CLUSTERS		
01	02	03
74.0 %	12.0 %	3.1 %
Data visualization and programming	Business Intelligence Data Driven Decision	Cloud Computing with DBMS
04	05	06
43.1 %	34.9 %	71.1 %
Big Data Java Hadoop Hive	Statistical Methods Business Intelligence	Stats Data modelling and Financial Planing
07	08	09
24.0 %	10.6 %	58.0 %
Cloud computing and Data Engineering	Probability and Applied Machine Learning	AI and Deep Learning and Natural Language Proc

# Part 2: Masters Course Sequencing and Design Ideology

CLUSTERS		
01	02	03
74.0 %	12.0 %	3.1 %
Data visualization and programming	Business Intelligence Data Driven Decision	Cloud Computing with DBMS
04	05	06
43.1 %	34.9 %	71.1 %
Big Data Java Hadoop Hive	Statistical Methods Business Intelligence	Stats Data modelling and Financial Planning
07	08	09
24.0 %	10.6 %	58.0 %
Cloud computing and Data Engineering	Probability and Applied Machine Learning	AI and Deep Learning and Natural Language Proc



## #4 Adding Coop / Internship –

The internship would help to gain the soft skills and industrial experience so to be prepared for full time jobs

## #3 Giving option of specialization -

A masters degree with a specialization certificate in skills like cloud computing can be of a great asset in job market

## #2 Course sequencing based on agglomerative learning –

The courses are sequenced in a way to build up knowledge step by step, so that a student gets full understanding of the subject matter

## #1 Course Evaluation based on % skills presence in jobs –

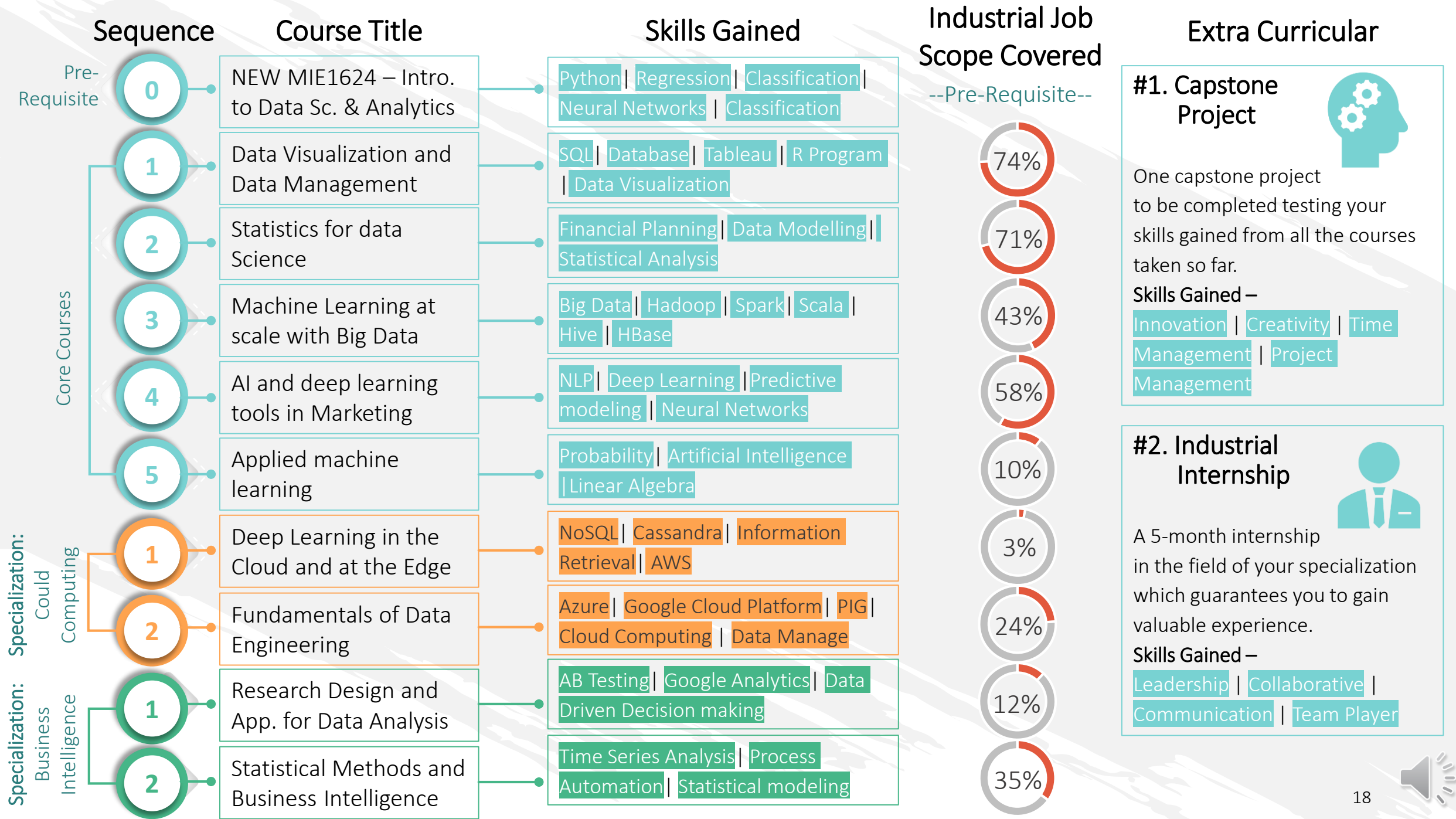
the course which have maximum presence in jobs are made mandatory core courses. (based on the skill clusters)





# Part 2 Result: Masters Program Design – *Master of Data Science and Artificial Intelligence*





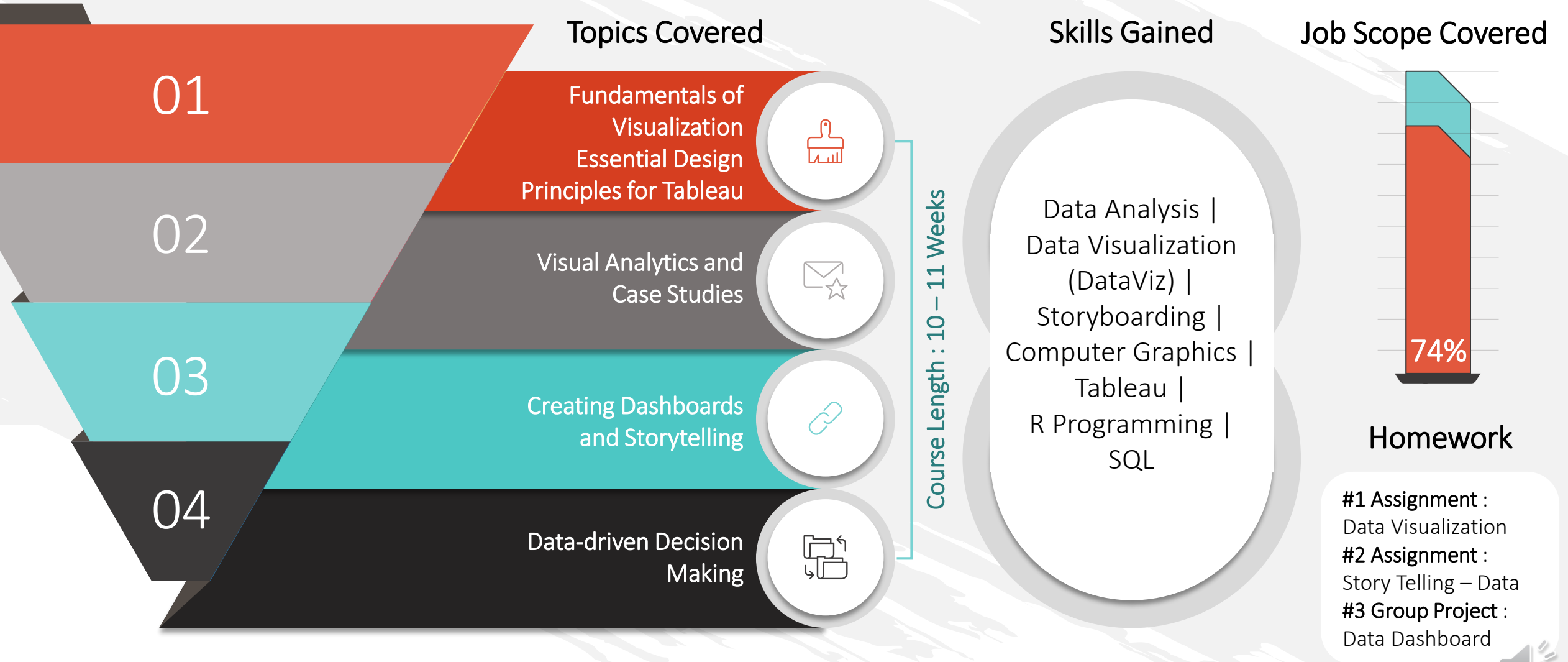
# Part 2 Result:

## *Master of Data Science and Artificial Intelligence*

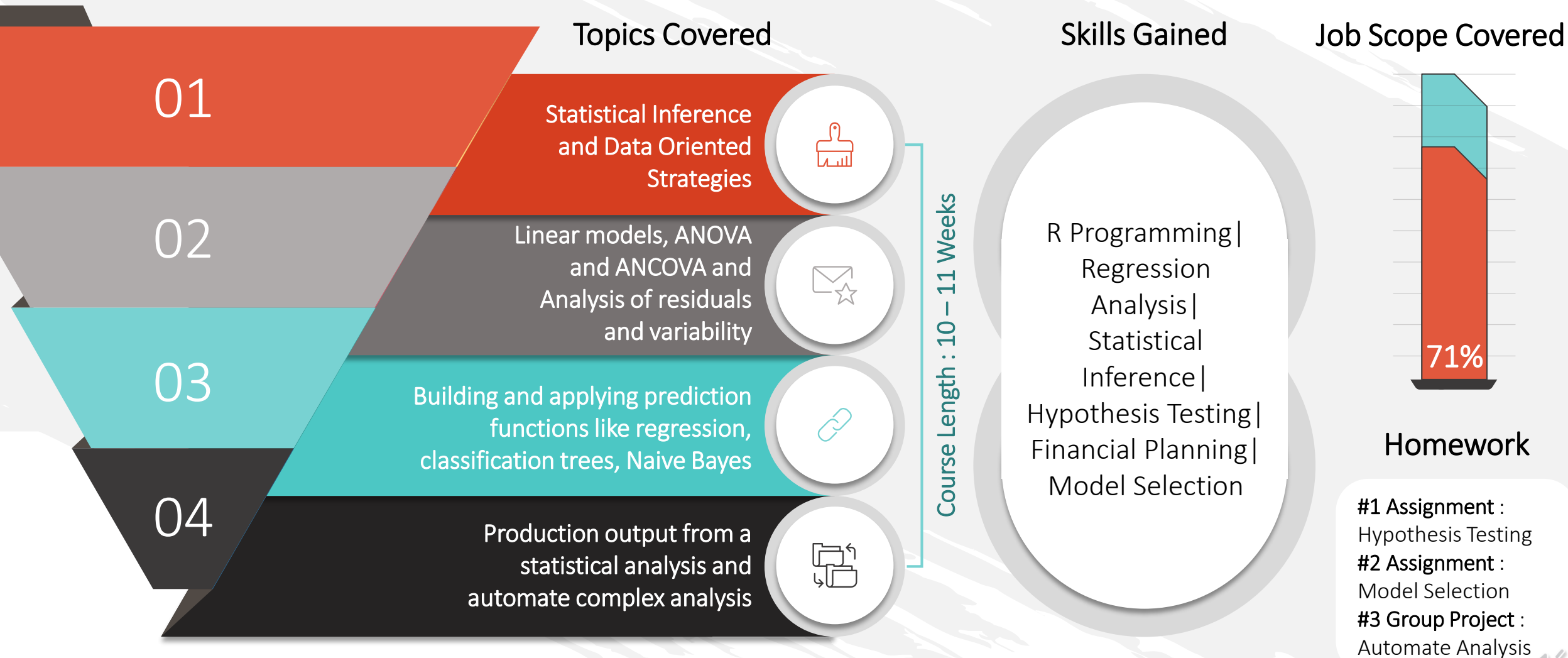
### Core Courses



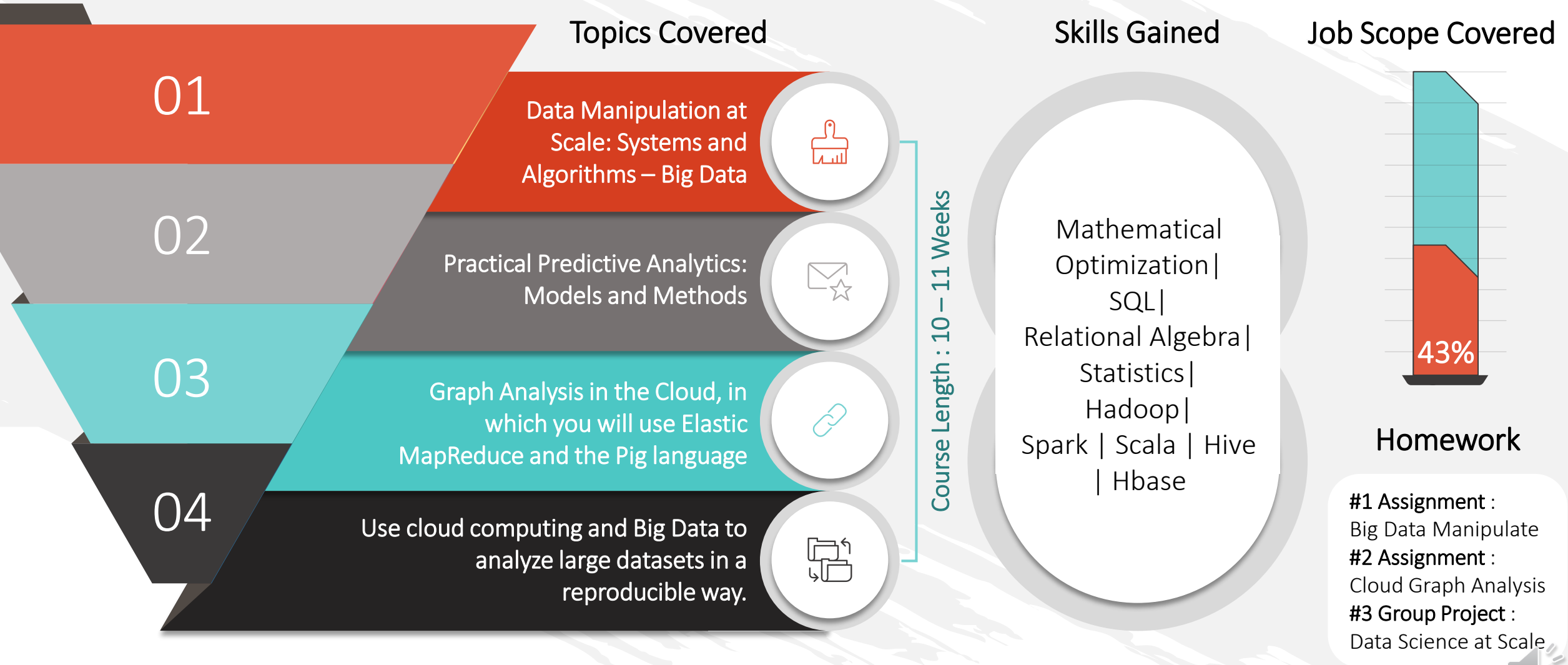
# Course 1: Data Visualization and Data Management



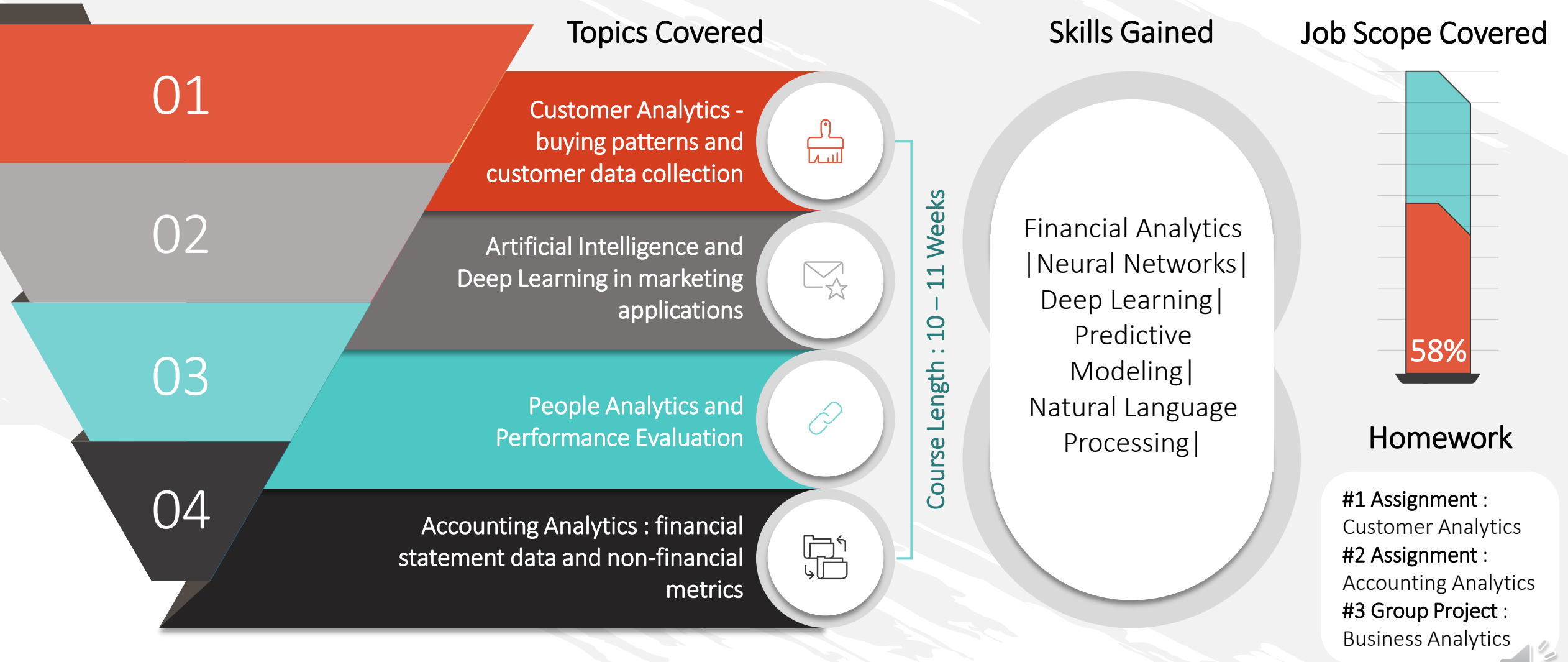
# Course 2: Statistics for Data Science



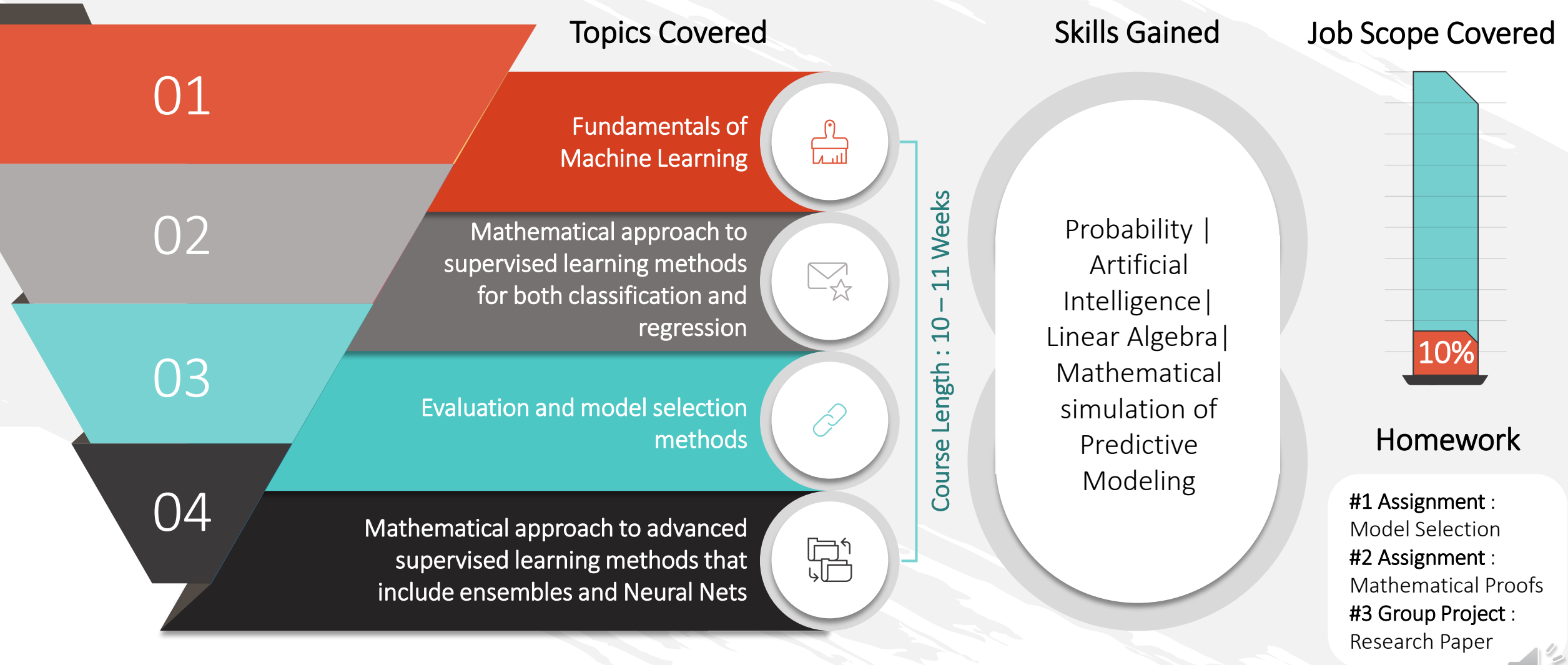
# Course 3: Machine Learning at Scale with Big Data



# Course 4: AI and Deep Learning tools in Marketing



# Course 5: Applied machine learning





# Part 3:

## Engineering Analytics Emphasis Course Selection Utility Recommender System



# Course Selection Utility

With the tool, the best courses are recommended to get the skills for a desired job

	Transportation Planner with Entuitive	Information Systems Analyst with Humber River Hospital	Data Scientist I with TD Bank
Required Course	APS1070: Foundations of Data Science and ML	APS1070: Foundations of Data Science and ML	APS1070: Foundations of Data Science and ML
Core Course	MIE1624: Introduction to Data Science and Analytics	MIE1624: Introduction to Data Science and Analytics	MIE1624: Introduction to Data Science and Analytics
Elective 1	CIV1507: Public Transport	MIE1623: Introduction to Healthcare Engineering	MIE1628: Big Data Science
Elective 2	APS1070: Supply Chain Mgmt and Logistics	APS502: Financial Engineering	APS1022: Financial Engineering II
Elective 3	CIV1532: Fundamentals of ITS and Traffic Mgmt	APS1005: Operations Research for Eng Mgmt	APS1040: Quality Control For Engineering Management



Thank You

