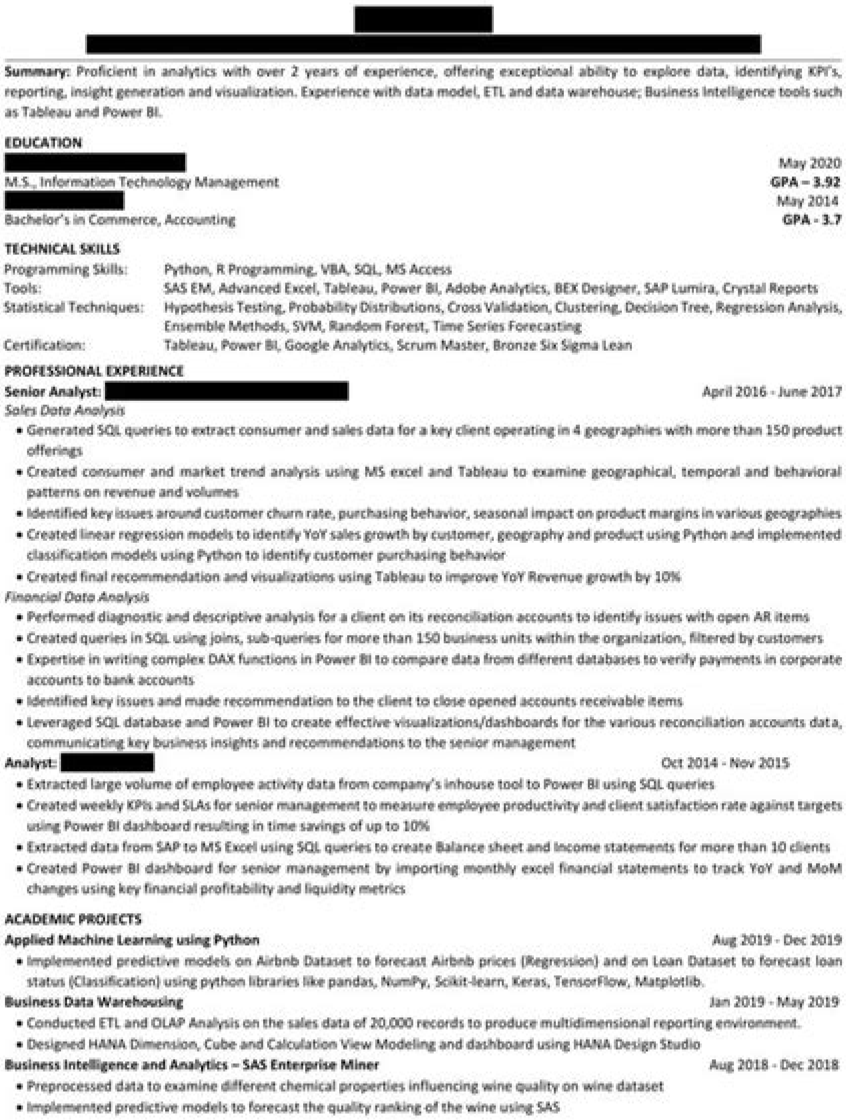
# Ibm Entry Level Data Scientist Coding Assessment



[redacted]

[redacted]

**Summary:** Proficient in analytics with over 2 years of experience, offering exceptional ability to explore data, identifying KPI's, reporting, insight generation and visualization. Experience with data model, ETL and data warehouse; Business Intelligence tools such as Tableau and Power BI.

**EDUCATION**

[redacted] May 2020
M.S., Information Technology Management **GPA – 3.92**
[redacted] May 2014
Bachelor's in Commerce, Accounting **GPA - 3.7**

**TECHNICAL SKILLS**

| | |
|---|---|
| Programming Skills: | Python, R Programming, VBA, SQL, MS Access |
| Tools: | SAS EM, Advanced Excel, Tableau, Power BI, Adobe Analytics, BEX Designer, SAP Lumira, Crystal Reports |
| Statistical Techniques: | Hypothesis Testing, Probability Distributions, Cross Validation, Clustering, Decision Tree, Regression Analysis, Ensemble Methods, SVM, Random Forest, Time Series Forecasting |
| Certification: | Tableau, Power BI, Google Analytics, Scrum Master, Bronze Six Sigma Lean |

**PROFESSIONAL EXPERIENCE**

**Senior Analyst:** [redacted] April 2016 - June 2017

*Sales Data Analysis*

- Generated SQL queries to extract consumer and sales data for a key client operating in 4 geographies with more than 150 product offerings
- Created consumer and market trend analysis using MS excel and Tableau to examine geographical, temporal and behavioral patterns on revenue and volumes
- Identified key issues around customer churn rate, purchasing behavior, seasonal impact on product margins in various geographies
- Created linear regression models to identify YoY sales growth by customer, geography and product using Python and implemented classification models using Python to identify customer purchasing behavior
- Created final recommendation and visualizations using Tableau to improve YoY Revenue growth by 10%

*Financial Data Analysis*

- Performed diagnostic and descriptive analysis for a client on its reconciliation accounts to identify issues with open AR items
- Created queries in SQL using joins, sub-queries for more than 150 business units within the organization, filtered by customers
- Expertise in writing complex DAX functions in Power BI to compare data from different databases to verify payments in corporate accounts to bank accounts
- Identified key issues and made recommendation to the client to close opened accounts receivable items
- Leveraged SQL database and Power BI to create effective visualizations/dashboards for the various reconciliation accounts data, communicating key business insights and recommendations to the senior management

**Analyst:** [redacted] Oct 2014 - Nov 2015

- Extracted large volume of employee activity data from company's inhouse tool to Power BI using SQL queries
- Created weekly KPIs and SLAs for senior management to measure employee productivity and client satisfaction rate against targets using Power BI dashboard resulting in time savings of up to 10%
- Extracted data from SAP to MS Excel using SQL queries to create Balance sheet and Income statements for more than 10 clients
- Created Power BI dashboard for senior management by importing monthly excel financial statements to track YoY and MoM changes using key financial profitability and liquidity metrics

**ACADEMIC PROJECTS**

**Applied Machine Learning using Python** Aug 2019 - Dec 2019

- Implemented predictive models on Airbnb Dataset to forecast Airbnb prices (Regression) and on Loan Dataset to forecast loan status (Classification) using python libraries like pandas, NumPy, Scikit-learn, Keras, TensorFlow, Matplotlib.

**Business Data Warehousing** Jan 2019 - May 2019

- Conducted ETL and OLAP Analysis on the sales data of 20,000 records to produce multidimensional reporting environment.
- Designed HANA Dimension, Cube and Calculation View Modeling and dashboard using HANA Design Studio

**Business Intelligence and Analytics – SAS Enterprise Miner** Aug 2018 - Dec 2018

- Preprocessed data to examine different chemical properties influencing wine quality on wine dataset
- Implemented predictive models to forecast the quality ranking of the wine using SAS

**IBM Entry Level Data Scientist Coding Assessment** is a pivotal step for aspiring data scientists seeking to join one of the largest technology companies in the world. This assessment serves as a gateway for candidates to demonstrate their programming skills, analytical thinking, and familiarity with data science concepts. As the demand for data-driven decision-making increases, IBM is keen on attracting individuals who can contribute effectively to their data-driven projects. In this article, we will explore the various aspects of the IBM entry-level data scientist coding assessment, including its structure, preparation strategies, and types of questions you might encounter.

Overview of the IBM Entry Level Data Scientist Role

Before delving into the specifics of the coding assessment, it is essential to understand the role of a data scientist at IBM. Data scientists at IBM are responsible for:

- Analyzing large data sets to derive actionable insights.
- Building predictive models using machine learning algorithms.
- Communicating findings to stakeholders through data visualization.
- Collaborating with cross-functional teams to enhance data-driven strategies.

Given these responsibilities, the coding assessment aims to evaluate candidates on their technical skills, problem-solving abilities, and understanding of statistical concepts.

## Structure of the Coding Assessment

The IBM entry-level data scientist coding assessment typically consists of the following components:

### 1. Online Assessment

The initial phase of the assessment is usually conducted online and may include:

- Multiple-choice questions: These questions assess your theoretical knowledge of data science concepts, programming languages, and statistics.
- Coding challenges: You will be required to solve programming problems using languages such as Python, R, or SQL.

### 2. Technical Interview

Candidates who perform well in the online assessment may be invited for a technical interview. This phase usually involves:

- Live coding exercises: Interviewers may ask you to solve problems in real-time, allowing them to observe your thought process and coding style.
- Behavioral questions: These questions assess your interpersonal skills and how you work in a team environment.

### 3. Case Study Presentation

In some instances, candidates may be required to complete a case study presentation where they analyze a dataset and present their findings. This tests not only your analytical skills but also your ability to communicate complex ideas effectively.

## Key Topics Covered in the Assessment

To excel in the IBM entry-level data scientist coding assessment, candidates should be well-versed in the following key topics:

### 1. Programming Skills

Proficiency in programming languages is crucial. The assessment may cover:

- Python: Familiarity with libraries such as NumPy, Pandas, and Scikit-learn for data manipulation and machine learning.
- R: Understanding of statistical modeling and data visualization techniques using ggplot2 and dplyr.
- SQL: Ability to write queries to extract and manipulate data from relational databases.

### 2. Statistics and Probability

A solid grasp of statistical concepts is essential for data analysis. Candidates should know:

- Descriptive statistics (mean, median, mode, variance).
- Inferential statistics (hypothesis testing, p-values, confidence intervals).
- Probability distributions (normal distribution, binomial distribution).

### 3. Data Wrangling and Preprocessing

Data often comes in raw forms that require cleaning and transformation. Essential skills include:

- Handling missing values.
- Normalizing and standardizing data.
- Encoding categorical variables.

### 4. Machine Learning Fundamentals

Understanding the basics of machine learning algorithms is critical. Candidates should review:

- Supervised vs. unsupervised learning.
- Common algorithms (linear regression, decision trees, clustering).
- Model evaluation metrics (accuracy, precision, recall, F1 score).

### 5. Data Visualization

Being able to visualize data effectively is key for communicating insights. Familiarity with tools such as:

- Matplotlib and Seaborn in Python.
- ggplot2 in R.
- Understanding the principles of effective data visualization.

## Preparation Strategies

Preparing for the IBM entry-level data scientist coding assessment requires a strategic approach. Here are some effective strategies:

### 1. Review the Basics

Ensure that you have a strong foundation in the core topics listed above. Utilize resources such as:

- Online courses (Coursera, edX, Udacity).
- Textbooks and reference materials.
- Video lectures on platforms like YouTube.

### 2. Practice Coding Challenges

Familiarize yourself with coding challenges by using platforms such as:

- LeetCode: Offers a wide range of coding problems that can help improve your algorithmic thinking.

- HackerRank: Provides specific challenges related to data science and machine learning.
- Kaggle: Participate in competitions and work on datasets to gain practical experience.

3. Mock Interviews

Conduct mock interviews with peers or mentors to practice live coding and behavioral questions. Consider using platforms like:

- Pramp: A peer-to-peer interview practice platform.
- Interviewing.io: Offers mock technical interviews with industry professionals.

4. Build a Portfolio

Creating a portfolio of projects can showcase your skills and knowledge to potential employers. Consider including:

- Data analysis projects using real-world datasets.
- Machine learning models with clear documentation.
- Visualizations that effectively communicate your findings.

5. Stay Updated on Industry Trends

The field of data science is continually evolving. Stay informed about the latest trends, tools, and technologies by:

- Following influential data scientists on social media.
- Reading blogs and articles related to data science.
- Attending webinars and conferences.

Conclusion

The IBM entry-level data scientist coding assessment is a critical step for candidates aspiring to join a leading technology company. By understanding the structure, key topics, and preparation strategies, you can significantly enhance your chances of success. Remember, the assessment is not just a test of your technical skills; it is also an opportunity to demonstrate your passion for data science and your ability to solve real-world problems. With dedication and thorough preparation, you can position yourself as a strong candidate for the role and embark on an exciting career in data science at IBM.

# Frequently Asked Questions

## What programming languages are typically assessed in the IBM entry-level data scientist coding assessment?

The assessment usually focuses on Python and SQL, as they are widely used in data science for data manipulation and analysis.

## What types of coding problems can I expect in the IBM entry-level data scientist assessment?

Candidates can expect problems involving data manipulation, statistical analysis, machine learning algorithms, and basic data visualization tasks.

## How should I prepare for the coding portion of the IBM entry-level data scientist assessment?

It's advisable to practice coding challenges on platforms like LeetCode or HackerRank, review data science concepts, and work on real datasets to enhance your skills.

## Is there a specific framework or library that I should be familiar with for the coding assessment?

Familiarity with libraries like Pandas, NumPy, and Scikit-learn in Python is beneficial, as they are commonly used for data manipulation and analysis.

## What is the format of the IBM entry-level data scientist coding assessment?

The assessment typically consists of a mix of multiple-choice questions and coding tasks that need to be solved in a timed environment.

## Are there any resources recommended for studying for the IBM data scientist coding assessment?

Yes, resources such as online courses on platforms like Coursera or edX, as well as books on data science and Python programming, are highly recommended.

## How much time is usually allotted for the coding assessment?

Candidates are generally given 60 to 90 minutes to complete the coding assessment, depending on the specific requirements of the test.

## Will I be penalized for incorrect answers in the IBM entry-level data scientist assessment?

Typically, there are no penalties for incorrect answers, but it's best to check the specific guidelines provided for the assessment.

## What skills are most important to demonstrate during the coding assessment?

Key skills include problem-solving, proficiency in programming, understanding of data structures, and the ability to apply statistical methods effectively.

## Can I use an online compiler during the IBM coding assessment?

Usually, the assessment platform will provide its own integrated development environment (IDE), and candidates are expected to use that for coding.

Find other PDF article:
https://soc.up.edu.ph/29-scan/files?docid=iXB89-4952&title=how-many-questions-are-on-the-texas-cdl-permit-test.pdf

# Ibm Entry Level Data Scientist Coding Assessment

**IBM**到底是做什么的？到底是以什么方式赚钱的？ **-** 知乎
Dec 28, 2013 · IBM到底是做什么的？到底是以什么方式赚钱的？以前认识IBM是在联想收购了IBM的PC部门（当时还以为是收购了IBM…），然后慢慢的知道IBM在服务器存储方面很牛（ …

应届生进入**ibm**怎么样？ **-** 知乎
IBM涨薪确实就是慢，已经比不上很多国内的民企了。我三年半只加了两次薪，当年猎头给我打电话还说每年有普调，每季度有机会调，科科，说白了普调也是一线经理可以不给你调的。我 …

如何高大上地使用 ThinkPad 的「小红帽」（TrackPoint）？ - 知乎
Nov 11, 2014 · 目前我选择thinkpad唯一的理由就是小红点。不要说苹果触摸板多好用，我用了一年多还是怀念小红点，最后换回了thinkpad。小红点上手确实有些难度，初次看见不直观不知 …

IBM和四大的咨询服务有什么差别？ - 知乎
自IBM于2002年并购普华永道，IBM Consulting的积累已经超过20年。IBM在2021年10月正式宣布IBM Consulting，其前称为IBM全球企业咨询服务。随后，2021年11月IBM公司正式拆分，– …

**IBM** 的 **POWER** 处理器的架构比 **X86** 强吗？ **-** 知乎
IBM 的 POWER 处理器的架构比 X86 强吗？据统计，2013年IBM服务器业务总营收为154亿美元，其中X86服务器业务营收为49亿美元，也就是说IBM的POWER营收是X86的两倍左右。

战略与执行-IBM业务领导力模型（BLM）
Nov 3, 2022 · BLM模型 (Business Leadership Model),中文名称叫做 业务领导力模型,也译为业务领先模型。在咨询行业,IBM的BLM模型,可以和著名的波士顿矩阵、SWOT分析以及迈克波特的 …

北京动态I|动态科技|成都ibmbjb|北京ibm二手笔记本水货|动态老袁
ThinkPad笔记本电脑,港货笔记本电脑,ThinkPad T14,Thinkpad X1 carbon,ThinkPad工作站,动态老袁,成都海滨,ThinkPad二手笔记本,苹果笔记本,苹果iPad,在北京，成都均有分店

*mac*如何成功安装*spss*，谢谢？ - 知乎
IBM SPSS Statistics是一款专业的 数据统计分析软件，无论是在学术研究领域，还是商业经营领域，都起着举足轻重的作用。那么 IBM SPSS Statistics 的mac版本应该如何安装呢。安装包在 …

*IBM*是一家怎样的公司？它的历史、现在及未来。 - 知乎

4、客户第一还是员工第一？老沃森时代，IBM被视作能够给予员工充分尊重的企业---不论是文化上，还是在利益上。但在现在IBM高层的语境里，客户和投资者往往是排位更加靠前的，在整 ...

为什么ibm的小型机大型机那么稳定? - 知乎

IBM小型机好多提高稳定性的技术十几二十几年前就有了，x86不是说没有，学不会，而是觉得没必要，为了可靠性提高一个9，成本提高10倍?

IBM到底是做什么的? 到底是以什么方式赚钱的? - 知乎

Dec 28, 2013 · IBM到底是做什么的？到底是以什么方式赚钱的？以前认识IBM是在联想收购了IBM的PC部门（当时还以为是收购了IBM...），然后慢慢的知道IBM在服务器存储方面很牛（ ...

应届生进入**ibm**怎么样？ **-** 知乎

IBM涨薪确实就是慢，已经比不上很多国内的民企了。我三年半只加了两次薪，当年猎头给我打电话还说每年有普调，每季度有机会调，科科，说白了普调也是一线经理可以不给你调的。我 ...

如何高大上地使用 *ThinkPad* 的「小红帽」（*TrackPoint*）？ - 知乎

Nov 11, 2014 · 目前我选择thinkpad唯一的理由就是小红点。不要说苹果触摸板多好用，我用了一年多还是怀念小红点，最后换回了thinkpad。小红点上手确实有些难度，初次看见不直观不知 ...

IBM和四大的咨询服务有什么差别? - 知乎

自IBM于2002年并购普华永道，IBM Consulting的积累已经超过20年。IBM在2021年10月正式宣布IBM Consulting，其前称为IBM全球企业咨询服务。随后，2021年11月IBM公司正式拆分，– ...

*IBM* 的 *POWER* 处理器的架构比 *X86* 强吗？ - 知乎

IBM 的 POWER 处理器的架构比 X86 强吗？据统计，2013年IBM服务器业务总营收为154亿美元，其中X86服务器业务营收为49亿美元，也就是说IBM的POWER营收是X86的两倍左右。

战略与执行**-IBM**业务领导力模型（**BLM**）

Nov 3, 2022 · BLM模型 (Business Leadership Model),中文名称叫做 业务领导力模型,也译为业务领先模型。在咨询行业,IBM的BLM模型,可以和著名的波士顿矩阵、SWOT分析以及迈克波特的 ...

北京动态*I*|动态科技|成都*ibmbjb*|北京*ibm*二手笔记本水货|动态老袁

ThinkPad笔记本电脑,港货笔记本电脑,ThinkPad T14,Thinkpad X1 carbon,ThinkPad工作站,动态老袁,成都海滨,ThinkPad二手笔记本,苹果笔记本,苹果iPad,在北京，成都均有分店

mac如何成功安装spss，谢谢? - 知乎

IBM SPSS Statistics是一款专业的 数据统计分析软件，无论是在学术研究领域，还是商业经营领域，都起着举足轻重的作用。那么 IBM SPSS Statistics 的mac版本应该如何安装呢。安装包在 ...

*IBM*是一家怎样的公司? 它的历史、现在及未来。 - 知乎

4、客户第一还是员工第一？老沃森时代，IBM被视作能够给予员工充分尊重的企业---不论是文化上，还是在利益上。但在现在IBM高层的语境里，客户和投资者往往是排位更加靠前的，在整 ...

为什么ibm的小型机大型机那么稳定? - 知乎

IBM小型机好多提高稳定性的技术十几二十几年前就有了，x86不是说没有，学不会，而是觉得没必要，为了可靠性提高一个9，成本提高10倍?

Prepare for the IBM entry level data scientist coding assessment with our essential guide. Discover tips

Back to Home