

Building an Application is as simple as “ABC” in IBM Cloud Pak for Data

—
Sanjit Chakraborty
Lucy Obus

Data and AI Forum 2019

Data and AI Forum / © 2019 IBM Corporation



Welcome to the session

Sanjit Chakraborty
Client Experience Team Lead



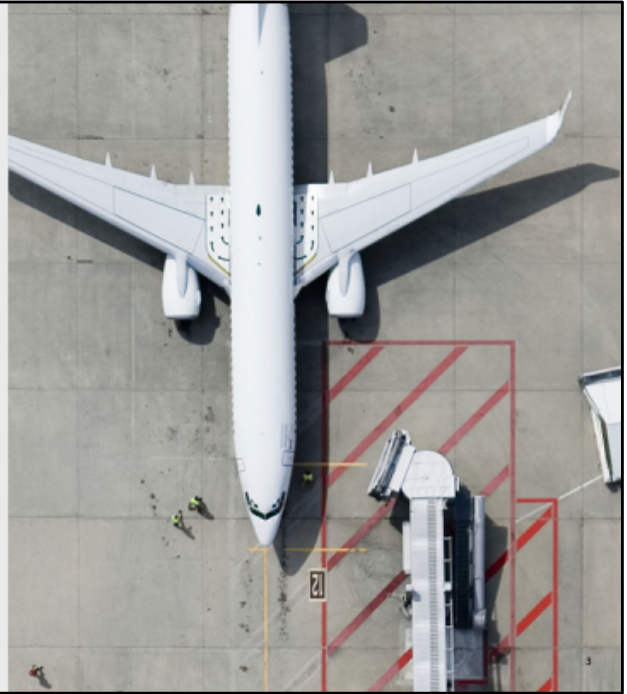
Lucy Obus
Design and User Experience Researcher



Let's get started

Travelbid

Data and AI Forum / © 2019 IBM Corporation





The Used Case:

- TravelBid is fictitious startup company, located at New York City.
- They provides discounted hotel and rental car reservation service within the city.
- TravelBid has a bulk of hotel and rental car vendors, from where they fulfill travel reservation for their clients.

Requirements are....

Looking for a data & analytics platform that can allows them easily implement an e-commerce system

Where they can have access their different vendor databases, pertaining to hotel and rental car reservation.

There should be an application interface, which allows potential clients named their price for hotel rooms and car rentals

Depends on client's named price, system generates list of hotel and car matches within the prices.

Challenges

Business Challenges

Need a **single user platform** to incorporate traditional **database search based on clients input** details, along with some **mechanism to predict** other possible combination of hotel and rental car choice based on clients named price.

Data and Analytics Challenges

Accumulate Data

Data sits across different databases

- Db2
- Mongo

Take lot of time and effort to access and merge data

Classifying Data

Ensuring data conforms to regulatory policies is a manual and time consuming process

Merging/Cleaning/
Transforming the Data takes up bulk of time

Explore Data

No proper tools for analyze the data and application development

As a startup, TravelBid wants to keep their operating cost as low as possible. Their main business challenge is come up with a single user platform to incorporate traditional **database search based on clients input** details, along with some **mechanism to predict** other possible combination of hotel and rental car choice based on clients named price.

They have several challenges from data and analytics prospective.

- Data sources are scattered over different databases, which takes lot of time and effort to access and merge data.
- As a travel service provider, TravelBid operated their business in certain regulatory environment. They need to ensure all their data must governed by necessary regulatory policies.
- Managing, cleaning and transform the data takes lot of their time.
- Right tools for analyze the data and develop applications are not available in their current environment.

Solution

IBM Cloud Pak for Data

A data and analytics platform that **simplifies** and **unifies** - **collect, organize and analyze data** to speed time to value from AI.

Capabilities

Collect

Connect to existing
Databases

Data Virtualization

Data Catalog

Organize

Enterprise Search

Business Rules/Terms

Auto Discovery

Analyze

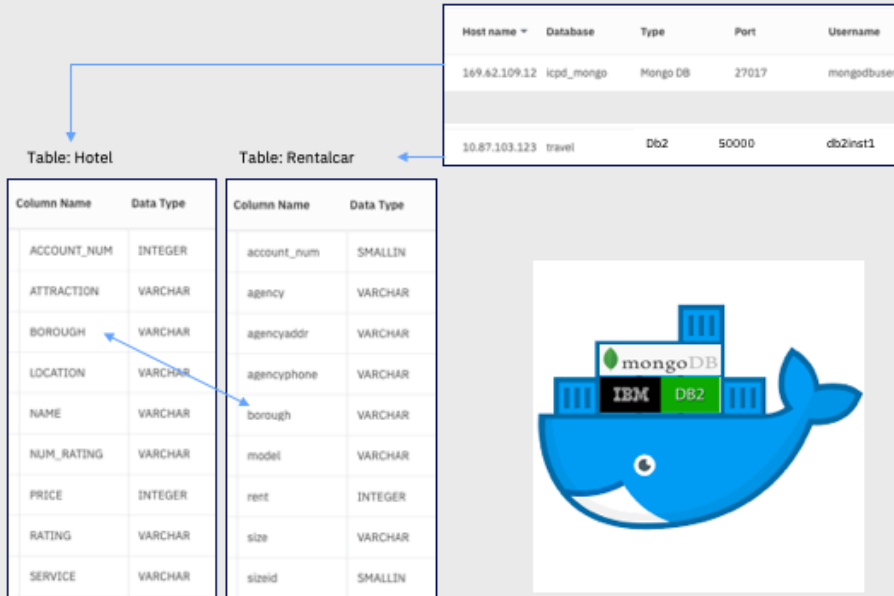
Open Source + Visual
tools

Published end-point

IBM Cloud Pak for Data is a possible platform that can fulfill all TravelBid's requirements.

Cloud Pak for data has three distinct pillars that simplifies and unifies - collect, organize and analyze data to speed time to value from AI.

Final Application: Data Sources



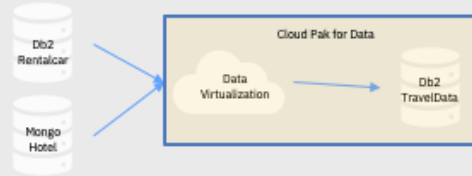
In next few slides you will see how we combined different aspects and built the application.

TravelBid largely used two data sources for their travel reservation business.

The HOTEL table resides on a Mongo database and stores information about hotels in New York City, where as RENTALCAR table is on Db2 database that store information about different rental car agencies and their fleets.

BOROUGH is the common primary key on these tables. It represents different sections within New York City.

Final Application: Virtual Table



```
CREATE VIEW TRAVELDATA#  
  AS SELECT "USER999"."RENTALCAR"."MODEL"  CARMODEL,  
            "USER999"."RENTALCAR"."AGENCY" CARAGENCY,  
            "USER999"."RENTALCAR"."SIZE"   CARSIZE,  
            "USER999"."RENTALCAR"."BOROUGH" BOROUGH,  
            "USER999"."RENTALCAR"."SIZEID" SIZEID,  
            "USER999"."HOTEL"."NAME"       HOTEL,  
            "USER999"."HOTEL"."RATING"    HOTELRATING,  
            ("USER999"."HOTEL"."PRICE" + "USER999"."RENTALCAR"."RENT") COST  
  FROM "USER999"."RENTALCAR", "USER999"."HOTEL"  
 WHERE "USER999"."RENTALCAR"."BOROUGH" = "USER999"."HOTEL"."BOROUGH";
```

You will use the Data Virtualization to merge two different data sources and build a virtual table (view).

Use this virtual table for analyzing data and run query.

Depend on the user name you used for login to the cluster, may need to change the schema name in this view statement. In this example, user name admin is associated with schema "USER999".

Final Application: Input

Recap of Requirements

- Build a data and analytics platform for rapidly develop an e-commerce system
- Potential **travelers named their price** for hotel rooms and car rentals
 - Price
 - Location
 - Hotel Rating
 - Car Type
- Depends on named price, system generates list of hotel and car matches within the prices

The screenshot shows a Shiny application interface with a dark theme. It has two main sections: 'Budget' and 'Hotels and Rental Cars'. The 'Budget' section contains a 'Price Range (\$)' slider with a range from 100 to 1,000 and a current value of 387. The 'Hotels and Rental Cars' section contains three dropdown menus: 'Borough' (set to 'Manhattan'), 'Hotel Rating' (set to '1'), and 'Car Type' (set to 'Economy'). A blue arrow points from the 'Price' requirement in the list to the 'Price Range' slider in the app.

If we recollect our requirements...

You need to build an application where potential clients named their price for hotel and car.

To get a better understanding and fulfill clients need accurately, you need to take some inputs from client., e.g., Price, Location, Hotel Rating, Car Type etc. So these inputs has to be part of your user interface, while building the R Shiny App.

Final Application: Project

The screenshot displays the Cloud Pak for Data interface. On the left is a navigation sidebar with options: Home, Projects, Connections, My instances, Collect, My data, Virtualized data, Organize, Analyze, Analytics dashboard, and Administer. The main area is divided into two panels. The top panel, titled 'Projects', shows a table of projects with columns for Name, Project Type, User Role, and Last Updated. The bottom panel, titled 'TravelRecommendation1', shows a list of assets under the 'Assets' tab, including Data sets, Notebooks, Scripts, Models, Model groups, and Analytics dashboards. A red arrow points from the 'Projects' section in the sidebar to the 'Projects' panel, and another red arrow points from the 'Assets' section in the sidebar to the 'Assets' panel.

Project
Collection of assets that use for data analysis

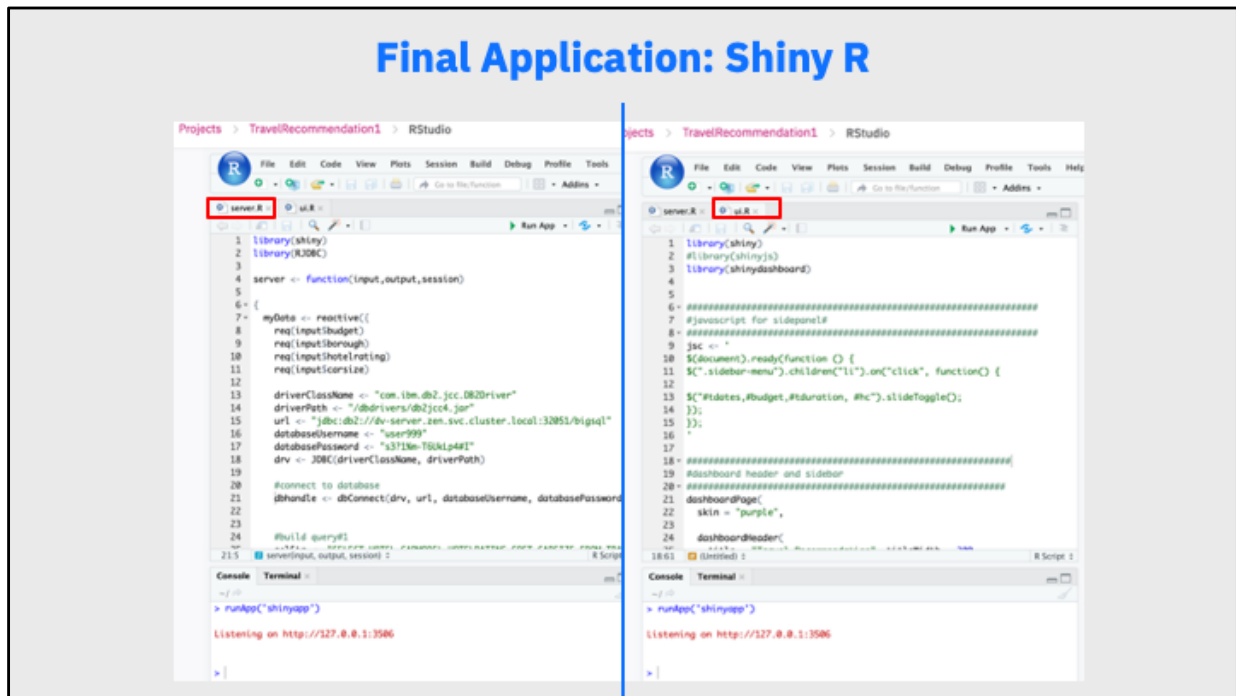
Assets

- Notebooks
- RStudio files
- Models
- Data assets (local files, data sources, and remote data sets)
- Scripts

In Cloud Pak for Data all required assets are clubbed under a project. Assets could be notebook, R scripts, data sets, models etc.

You create a project first to collect all asserts. Inside that project you create R Shiny application.

Final Application: Shiny R



As you know, R is an open source programming language that used for statistical computing and graphical representation of data. The R language is widely used among statisticians and data miners for developing statistical software and data analysis.

R shiny developed on R programming language that provide an easy web application framework for turn your data into an interactive web application.

The best part it simple and easy. You don't need to know whole lot of HTML or JavaScript. But you you can make your app more interesting by using some HTML.

Shiny app has two main components: 1) The user interface script to control the layout and appearance of application, 2) server script is use for connect and manipulate data.

There are tons of information on internet to learn about Shiny R. One of them is 'shiny.rstudio.com'.

Final Application: Web Interface

Input / Budget

- Price Range
- Borough
- Hotel Rating
- Car type

Output / Best Price

- Cost Effective

Output / Best Experience

- Unique Experience

Deploy Application

https://169.62.109.12:31843/dsx-publish-view-app/^^all_/rstudio/Travel_Recommendation

The screenshot shows a web interface titled "Travel Recommendation". On the left, there are input fields for "Price Range (\$)" with a slider, "Borough" (a dropdown menu), "Hotel Rating" (a dropdown menu), and "Car Type" (a dropdown menu). Below these is a "Submit" button. On the right, there are two tables. The top table is titled "Best Price" and lists ten recommendations with columns for HOTEL, CARMODEL, HOTELRATING, COST, and CARSIZE. The bottom table is titled "Best Experience" and lists ten recommendations with the same columns.

HOTEL	CARMODEL	HOTELRATING	COST	CARSIZE
Hotel 309	Dodge Avenger	3	160.00	Midsize
Hotel 309	Nissan Sentra	3	165.00	Midsize
Royal Park Hotel and Hostel	Dodge Avenger	3	170.00	Midsize
Hotel 309	Chevrolet Cruze	3	174.00	Midsize
Hotel 309	Dodge Avenger	3	174.00	Midsize
Royal Park Hotel and Hostel	Nissan Sentra	3	175.00	Midsize
Hotel 309	Toyota Camilla	3	177.00	Midsize
Hotel 309	Toyota Camilla	3	177.00	Midsize
Royal Park Hotel and Hostel	Chevrolet Cruze	3	184.00	Midsize
Royal Park Hotel and Hostel	Dodge Avenger	3	184.00	Midsize

HOTEL	CARMODEL	HOTELRATING	COST	CARSIZE
Harlem YMCA	Ford Mustang	3.5	350.00	Convertible
Hampton Inn Manhattan/Times Square South	Cadillac ATZ	4	350.00	Luxury
Colonial House Inn	Cadillac ATZ	4.5	350.00	Luxury
City Club Hotel	Lincoln MKS	4	350.00	Luxury
Hampton Inn Manhattan/Times Square South	Lincoln MKS	4	350.00	Luxury
City Club Hotel	Lincoln MKS	4	350.00	Luxury
Hampton Inn Manhattan/Times Square South	Lincoln MKS	4	350.00	Luxury
City Club Hotel	Cadillac ATZ	4	350.00	Luxury
Hampton Inn Manhattan/Times Square South	Cadillac ATZ	4	350.00	Luxury
City Club Hotel	Cadillac ATZ	4	350.00	Luxury

The application provides a Web based interface. It takes necessary information as input from client and depends on input information it suggests travel packages.

Left part used for get Input from clients.

- Client provides their daily budget for hotel and rental car for staying at New York City.
- Application also collect information about the place where they want to stay, their preferred hotel rating and a car type they want to use.

Based on the input, application generates two sets of reservation suggestions.

The top-right section shows, ten cost effective, or cheapest combination of of hotels and rental car price for a day, which based on client's hotel rating and car type.

Bottom section provides list of ten unique experience hotel and car combination within client's budget.

Cloud Pak for Data has an inbuilt web server under Watson Studio. Once the application ready you can deploy it on the platform and anyone with the published web link can easily access it on the internet. This application can be govern by the Cloud Pak for Data.

What we've done to get here:

Cloud Pak for
Data (CPD)
installed
and available

Relational
databases
(ex: Db2,
MongoDB)
created
and available

Business
glossary and
governance
policies created
and available for
import

Analytics project
created with R
Shiny

Access and sign in

Clusters

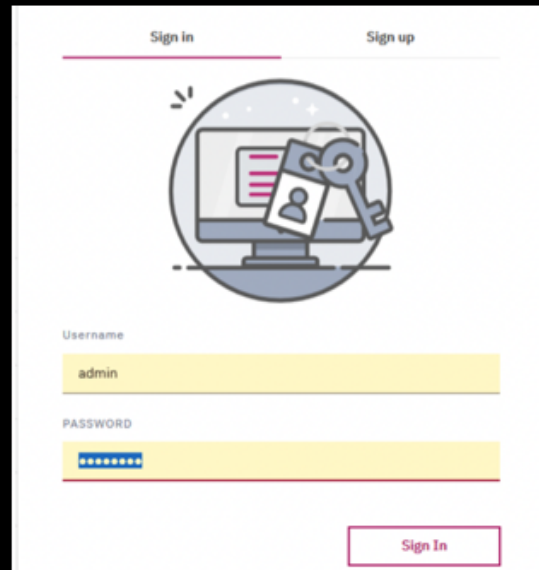
- <https://ibm.biz/296cluster1>
- <https://ibm.biz/296cluster2>
- <https://ibm.biz/296cluster3>
- <https://ibm.biz/296cluster4>
- <https://ibm.biz/296cluster5>

Users

- ctp1, ctp2, ctp3, ctp4, ctp5

Password

- ctpctp



The image shows a user interface for signing in or signing up. At the top, there are two tabs: "Sign in" (selected) and "Sign up". Below the tabs is a large circular icon containing a computer monitor, a key, and a user profile card. Underneath the icon, there are two input fields: "Username" with the text "admin" and "PASSWORD" with masked characters. A "Sign In" button is located at the bottom right of the form.

There are five Cloud Pak for Data clusters available for you to try out this lab. You need to follow the lab instruction to build the Shiny R application.

Instruction is available on your lab computer or you can download it from following address:
<https://drive.google.com/file/d/1cBBUjt19T7shINexjVM2i9KFwJ5A9wjn/view?usp=sharing>

But wait, there's more!

Additional questions?
Want to learn more about ICPD?

**Come visit us at the
Level 2 Atrium Foyer**
spoiler: there's swag



Don't miss out on:

Check the event app for schedule of **Expert Advice** sessions for topics on:

- Collect (Hybrid Data Management)
- Organize (Unified Governance & Integration)
- Analyze (Data Science & AI)
- Infuse (Business Analytics, Watson Apps & Solutions, Digital Business Automation)
- Modernize (IBM Cloud Pak for Data)

Giving **feedback** on the sessions through the **event app**.

Writing an IBM product review at the **Gartner Peer Insight booth**.

Visit community.ibm.com to join your peers and learn from experts across 12 different **IBM communities**.

Thank you

Sanjit Chakraborty

SanjitC@us.ibm.com



Lucy Obus

Lucy.Obus@us.ibm.com





Copyright © 2019 by International Business Machines Corporation (IBM). No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights – use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. This document is distributed "as is" without any warranty, either express or implied. In no event shall IBM be liable for any damage arising from the use of this information, including but not limited to, loss of data, business interruption, loss of profit or loss of opportunity. IBM products and services are warranted according to the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply."

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environment. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. IBM expressly disclaims all warranties, expressed or implied, including but not limited to, the implied warranties of merchantability and fitness for a particular purpose.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com, Aspera®, Bluemix, Blueworks Live, CICS, Clearcase, Cognos®, DOORS®, Eryptoria®, Enterprise Document Management System™, FASP®, FileNet®, Global Business Services®, Global Technology Services®, IBM ExperienceOne™, IBM SmartCloud®, IBM Social Business®, Information on Demand, ILGO, Medimo®, MQIntegrator®, MQSeries®, Netcool®, OMEGAMON, OpenPower, PureAnalytics™, PureApplication®, pureCluster™, PureCoverage®, PureData®, PureExperience®, PureFlex®, pureQuery®, pureScale®, PureSystems®, QRadat®, Rational®, Rhapsody®, Smarter Commerce®, SoDA, SPSS, Sterling Commerce®, Storwize®, Tivoli®, Trusteer®, Unica®, urbancode®, Watson, Websphere®, Worklight®, X-Force® and System z®/OS are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/entry/ucopyright.shtml

