

IBM CLOUD PACK FOR DATA

Datstage workshop!

DATASTAGE

- IBM® DataStage® is an industry-leading data integration tool that helps you design, develop and run jobs that move and transform data. At its core, the DataStage tool supports extract, transform and load (ETL) and extract, load and transform (ELT) patterns.

INTEGRATE DATA FROM MULTI CLOUD

- The story for the tutorial is that Golden Bank needs to adhere to a new regulation where it cannot lend to underqualified loan applicants. As a data engineer at Golden Bank, you currently use DataStage to aggregate your anonymized mortgage applications data with the mortgage applicants' personally identifiable information. Lenders use this information to help them decide whether they should approve or deny mortgage applications.
- Your leadership added some risk analysts who calculate daily what interest rate they recommend offering to borrowers in each **credit score** range. The spreadsheet includes credit score information for each applicant, the applicant's **total debt**, and an **interest-rate lookup table**. Lastly, load your data into a target output CSV file. The presentation follows the tutorial discussed here: [IBM cloud tutorial link](#)

TASKS

- Create the sample project "Multicloud data integration"
 - Task 1: Run an existing sample flow.
- Edit the existing datastage flow
 - Task 2: Specify a key column for the Join stage.
 - Task 3: Add credit score data from a PostgreSQL database.
 - Task 4: Add a Join stage to join the credit score data with the applicant and application data.
 - Task 5: Add a Transformer stage to calculate total debt.
 - Task 6: Add interest rate data from a MongoDB database.
 - Task 7: Add a Lookup stage to look up interest rates for applicants.

TASKS CONTD...

- Task 8: Edit the Sequential file node and run the DataStage flow.
- Task 9: Publish to Mortgage Approvers catalogue.

THANK YOU !

