Climate Change, Kiwibank and SAS Data Quality – What's the Connection?

Ian Munro, Risk Infrastructure Architect

The Journey So Far

- SAS Data Quality chosen in 2018 to deal with bank-wide data quality issues

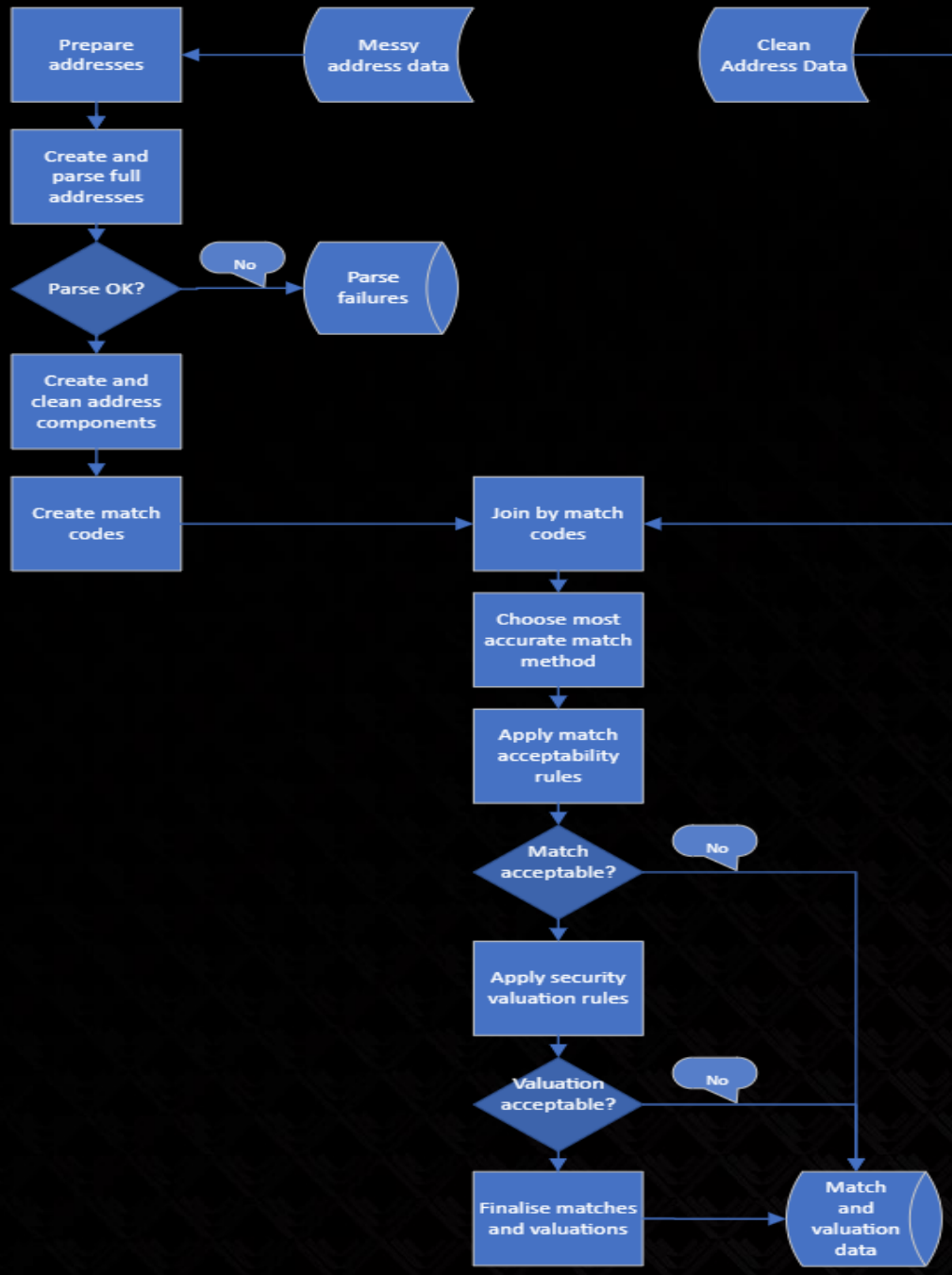
- Kiwibank Risk use case
  - Requirement - get current property security valuations
  - Challenge - join address sources with no common address key
  - Solution - match property security addresses to external address databases using SAS DQ

- External address databases each with own address key
  - NZ Post delivery addresses - DPID
  - Valocity property addresses and valuations - DIPID
  - Land Information NZ property addresses - address_id
  - LOQATE NZ addresses (supplied with SAS DQ) – AddressKey

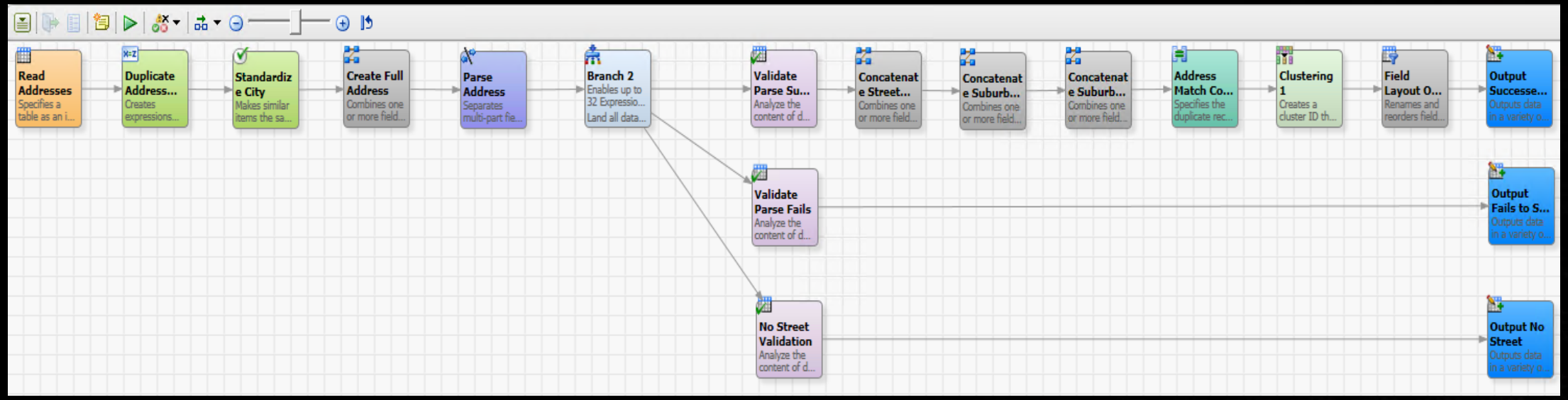
- Kiwibank property addresses with own security key
  - Addresses include street, suburb, city and postcode
  - Property valuation at time of application

SAS Data Quality Address Matching

- Bank address data has typical DQ issues – requires parsing
- External address data is cleaner but not error-free
- Create match codes for fuzzy matching
- Join addresses by match codes, choosing most accurate method
- Acceptably matched addresses have accurate address keys and geocoding
- Use keys to join to current valuation and now also climate change data



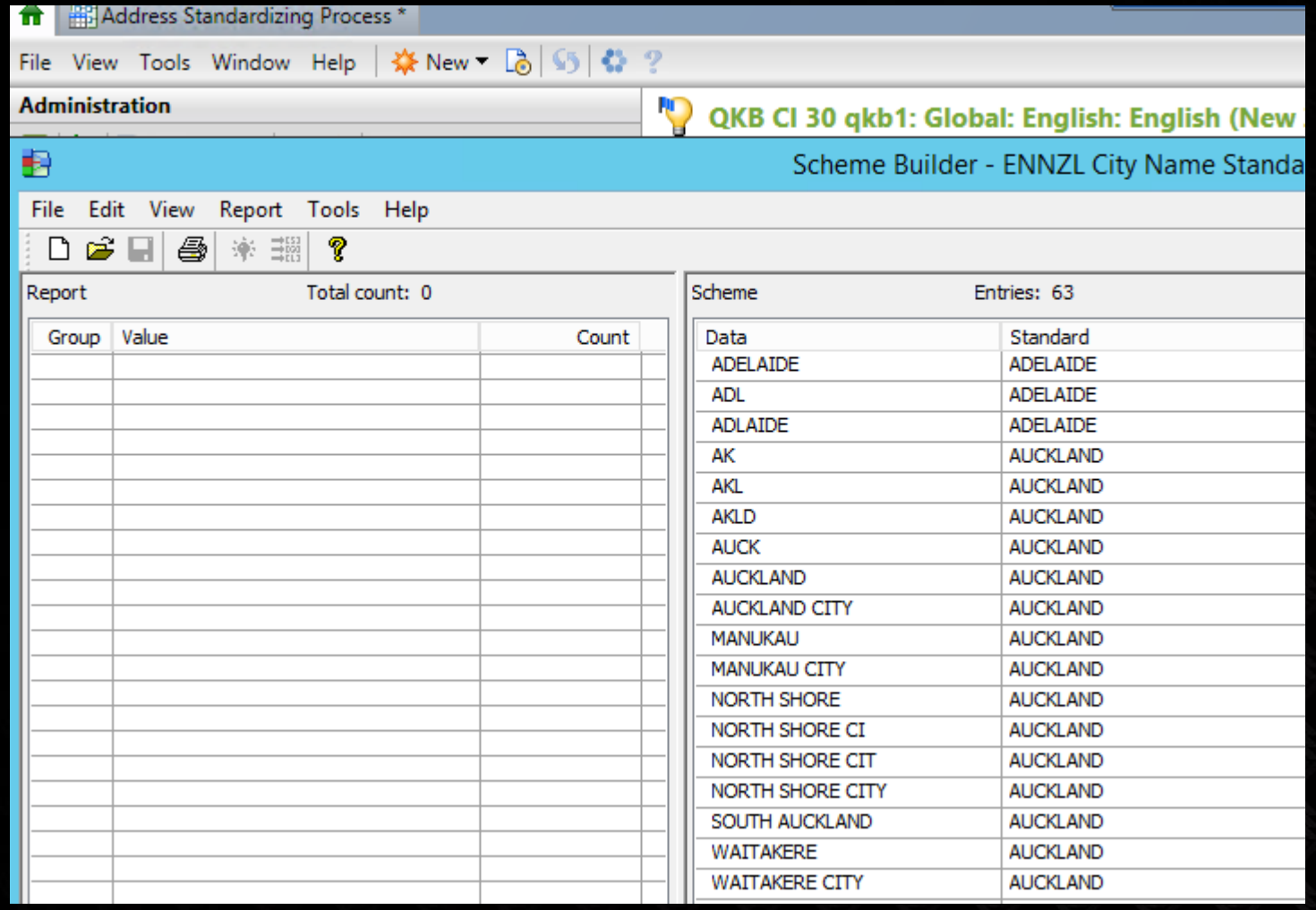
SAS DQ - DataFlux Data Management Studio



- Uses process flows like those in Enterprise Guide projects and SAS Studio flows
- This DQ data job implements the “messy” address process
- Once built, deploy jobs to SAS DQ server to run as scheduled processes

Quality Knowledge Base (QKB)

- SAS Data Quality contains a New Zealand-specific QKB
- Further customisations are usually required to improve accuracy
- The displayed example deals with the Auckland City amalgamation



SAS DQ Secret Sauce Examples

- Run DataFlux jobs in programs

```
data _null_;
    rc        = dmsrvauthdom("DMServerAuth");
   jobid      = dmsrvBatchJob('My Dataflux Job.ddf', 'http://DataFluxServer.MyCompany.com', 21036);
   jobstatus = dmsrvJobStatus(jobid, 'http://sasdq27.corp.bank.nzpfs.co.nz', 21036, -1, 5);
   put _all_;
run;
```

- Create DQ match codes in programs

```
%DQLOAD(DQLOCALE=(ENNZL), DQSETUPLOC='C:\ProgramData\SAS\SASQualityKnowledgeBase\CI\CI30_qkb1');

data want;
  set have;
  Street_Address_MatchCode = dqmatch(Street_Address, 'Address 5 Digit Extension - KB', 85,'ENNZL');
  Suburb_City_Postcode_MatchCode = dqmatch(Suburb_City_Postcode, 'City - State/Province - Postal Code', 85,'ENNZL');
  Suburb_City_MatchCode = dqmatch(Suburb_City, 'City - State/Province - Postal Code', 85,'ENNZL');
run;
```

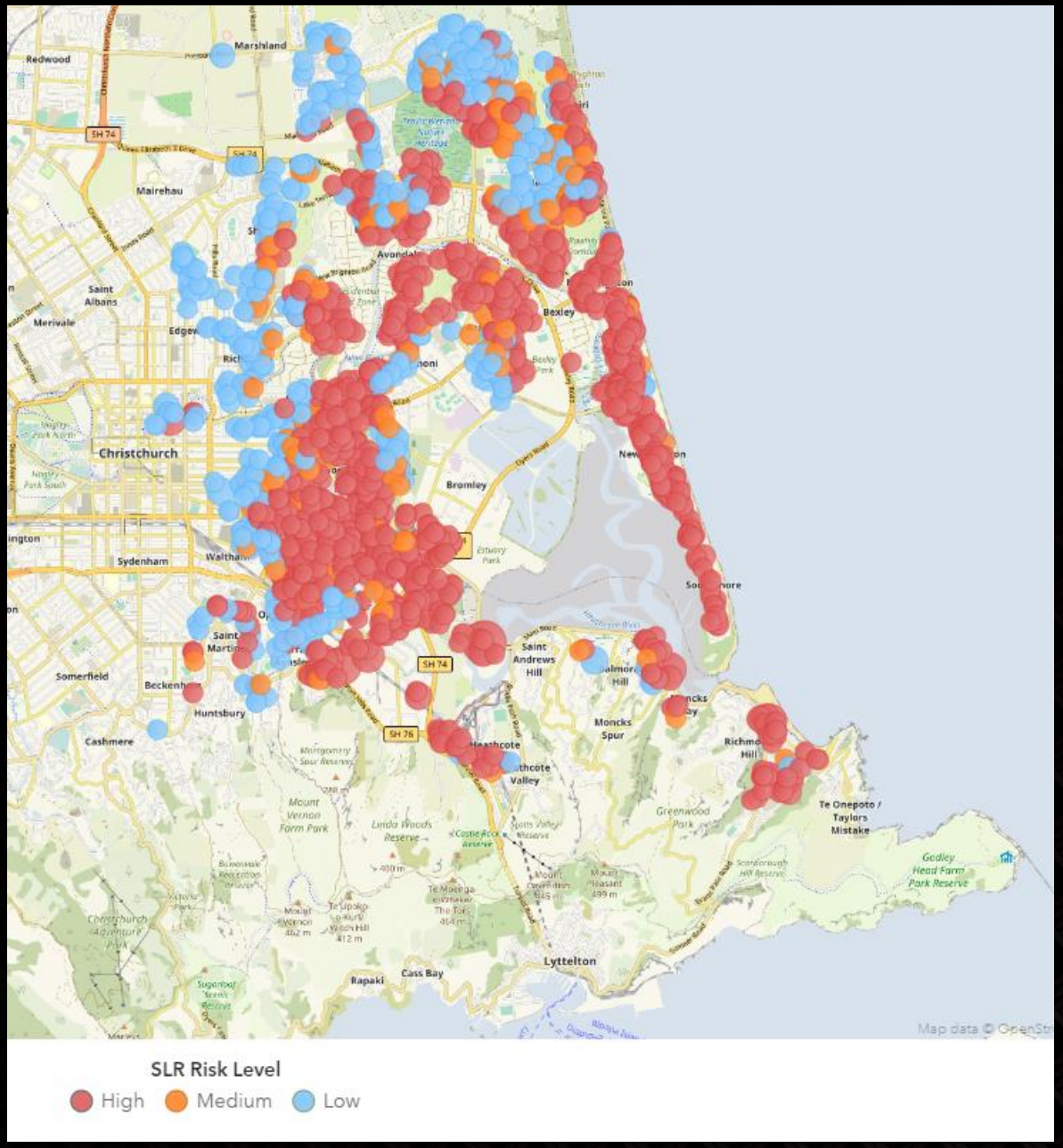
- SAS DQ is resource-intensive so we use SAS/CONNECT to integrate and parallel process DQ jobs

```
 * Start 4 SAS DQ sessions.;
 signon remote = SASDQ;
 %syslput _all_ / remote = SASDQ;

 %let SASDQ2 = &SASDQ;
 signon remote = SASDQ2;
 %syslput _all_ / remote = SASDQ2;

 %let SASDQ3 = &SASDQ;
 signon remote = SASDQ3;
 %syslput _all_ / remote = SASDQ3;

 %let SASDQ4 = &SASDQ;
 signon remote = SASDQ4;
 %syslput _all_ / remote = SASDQ4;
```

Climate Change Reporting – Kiwibank's Challenges

- Mandatory reporting for large NZ organisations
- Reserve Bank requires all NZ banks to report on climate change risks
- Climate change modelling is a very immature science and data is sparse
- SLR, pluvial and fluvial flooding are the first cabs off the rank

Sea Level Rise (SLR) Property Risks in Christchurch



Climate Change, Kiwibank and SAS Data Quality

- Climate change reporting introduces major new data challenges
  - Match and locate all property securities
  - External address keys are essential
  - Sourcing and joining climate change modelling data

- Kiwibank is meeting the challenges by:
  - Using SAS DQ to match to external address data sources and get external address keys
  - Working with providers of climate change modelling data to enhance and expand reporting capabilities