Novedea Systems Inc, 1750 N Collins Blvd Suite 212 Richardson, TX- 75080 Ph: 972-918-0900 Fax: 972-918-0901

Confidential –Do not distribute.

1003 Application Processing Requirements Document V7

ETL, Data ware & Analytics

10/22/2013

C:\Novedea\logo\logo.gif

Ram Kotamaraja

# Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version | Author | Notes |
| 5/22/2013 | Initial | Ram Kotamaraja | Created initial document version |
| 5/25/2013 | V1 | Ram Kotamaraja | Added project diagram |
| 5/26/2013 | V2 | Ram Kotamaraja | Added placeholders for questions to be answered and ftp info where users can upload their files |
| 6/07/2013 | V3 | Ram Kotamaraja | Added script task validation to story 2 and added questions to the stories |
| 6/29/2013 | V4 | Ali Muwwakkil | Added Story’s 6-10 |
| 7/15/2013 | V5 | Chaitu Sagam | Added Interview Questions in Appendix 2 |
| 10/.15/2013 | V6 | Chaitu Sagam | Updated Data types and Length field according to Excel Data Source in Appendix 1 |
| 10/23/2012 | V7 | Ram Kotamaraja | Updated business back ground information and removed interview questions from the stories. |
| 11/26/2015 | V8 | Ram Kotamaraja | Raefactored the project into multiple sprints and releases to accommodate more agile development. |

Table of Contents

[Revision History 1](#_Toc436286629)

[Business Background 3](#_Toc436286630)

[Project Information 3](#_Toc436286631)

[Project Diagram 4](#_Toc436286632)

[Project Methodology: 5](#_Toc436286633)

[Agile Methodology stories: 6](#_Toc436286634)

[R1:SP1: Story 1: Gather 1003 Mortgage application input data 6](#_Toc436286635)

[R1:SP2: Story 2: Read input data 7](#_Toc436286636)

[R1:SP2: Story 3: Insert data into staging tables 7](#_Toc436286637)

[R1:SP3: Story 4: Validate & Error Handling staging data 8](#_Toc436286638)

[R1:SP3: Story 5: Insert data into ODS database 8](#_Toc436286639)

[R1:SP4: Story 6: Insert data into OLAP / Data warehouse / Dimensional Model 9](#_Toc436286640)

[R2:SP5: Story 7: Generate Reports – Loans to Date 10](#_Toc436286641)

[R2:SP5: Story 8: Generate Reports – Loans to Date / Demographics 11](#_Toc436286642)

[R2:SP6: Story 9: Generate Reports – Loan Overview 12](#_Toc436286643)

[R2:SP6: Story 10: Report Functionality 13](#_Toc436286644)

[Appendix 1: Form Fields Information 15](#_Toc436286645)

[Appendix 2: Questions to explore as you do the project 16](#_Toc436286646)

[R1:SP1: Story 1: 16](#_Toc436286647)

[R1:SP2: Story 2: 16](#_Toc436286648)

[R1:SP2: Story 3: 16](#_Toc436286649)

[R1:SP3: Story 4: 17](#_Toc436286650)

[R1:SP3: Story 5: 17](#_Toc436286651)

[R1:SP4: Story 6: 17](#_Toc436286652)

[R2:SP5: Story 7, 8 & R2:SP6: 9,10: 18](#_Toc436286653)

[Additional Questions: 19](#_Toc436286654)

# Business Background

**Client:** Revolution Mortgage Group (RMG)

**About RMG:** RMG has built a strong reputation as an outstanding mortgage lender, serving the lending needs of individual homebuyers, real estate professionals and builders throughout their lending network.  With offices in New York, Boston, Raleigh, Dallas, Houston, Austin, Atlanta and Tampa, RMG is able to service customers in New York, Massachusetts, North Carolina, Texas and Florida.

RMG is a full service mortgage lender with an experienced mortgage loan offers offering expertise in every area of mortgage lending, from purchase to refinance to construction lending. RMG have access to a full range of mortgage sources and all of RMG lending specialists are dedicated to finding the right loan-with the best rates, terms and costs-to meet their customer’s unique needs.

As a subsidiary of North Beach Bank, Revolution Mortgage Group has the resources and support of a community bank with over 250 years of experience. RMG customers not only have access to the best loans available in the marketplace, but also can review loan alternatives, and even apply for loan, at their convenience, online - 24 hours a day.

# Project Information

**Project Description:** RMGreceives 1003 mortgage applications from customers. RMG receives application from multiple sources such as online, email/fax and from Zillow. An ETL solution should be developed to consolidate the loan applications that are received into one Loan Application Database. ETL process should validate the loan applications for all the required fields and for field format. If there are any problems with the applications, loan officers are notified about the errors. Loan officers will review the application errors, fix the errors and resubmits the application for further validation. Loan officers will review the valid applications on a daily basis and either approves the loan or rejects the loan. Both approved and rejected loans are then loaded into data warehouse for reporting purposes.

**Project Input data:** Applications are received in the following formats:

1. Online : Customers can visit RMG website and fill the form at: <https://www.secureloandocs.com/agreement.php?id=28788498>. The data captured using online application is stored in a OLTP database.
2. Email or Fax: Customers can fill Fannie Mae application and email or fax it directly to RMG loan service representative. Sample Fannie Mae application is located at <https://www.fanniemae.com/content/guide_form/1003rev.pdf>. Loan officers will manually fill this data into a spread sheet on a daily basis and sends it to FTP location on a daily basis.  This CSV file will have all the applications that are received during the day.
3. Zillow: Customers can go to [Zillow.com](http://Zillow.com), search for mortage rates and apply for mortage with RMG. Zillow will send the 1003 application in an XML format real time. This xml file is placed in an FTP location. This xml file always has one loan application.

**Data validations:** Data validation should be done according to the screen shots from <https://3185655371.secure-loancenter.com/WebApp/Start.aspx?>

**Data warehouse:** Data warehouse should be created to meet future growth of the group across USA.

**Reports:** Reports developed should provide intelligence to business on an hourly, daily, weekly, monthly, quarterly and yearly basis.

# Project Diagram



# Project Methodology:

RMG is very particular about getting high quality project deliverables. Agile methodology will be used to implement this project. Agile methodology will divide project into Releases, Sprints and User stories.

* A Project has multiple releases, with each release delivering a business value to its customers.
* A Release contain one or more sprints.
* A Sprint, with a time boxed scope, implement one or more stories in that sprint.
* A Story, an end user driven business function may contain one or more use cases. Often in the business, a story is interchangeably used with a use case (two terms from different generations). A story tends to be driven more by the business user perspective where as use cases used to represent the system perspective.

This mortgage project will use the following Agile methodology structure:

|  |  |  |
| --- | --- | --- |
| Release | Sprint | Story |
| Release 1 (R1) | Sprint 1 (SP1) | Story 1 |
| Sprint 2 (SP2) | Story 2  Story 3 |
| Sprint 3 (SP3) | Story 4  Story 5 |
| Sprint 4 (SP4) | Story 6 |
| Release 2 (R2) | Sprint 5 (SP5) | Story 7  Story 8 |
| Sprint 6 (SP6) | Story 9  Story 10 |

*Note: You are encouraged to allocate a max of one week of time for each sprint and complete the project with in six weeks’ or less time.*

# Agile Methodology stories:

## R1:SP1: Story 1: Gather 1003 Mortgage application input data

**Narrative:**

As a Developer

I want to gather multiple types of 1003 mortgage application input data from various sources

So that I can process the data using ETL

**Scenario 1 - Excel Data:** Get 1003 mortgage application data in excel format from Loan officers

Loan officers receive 1003 mortgage applications from borrowers in an email or fax or via phone. They will capture each application data into an excel file and upload that excel file to an ftp folder at the end of the day.

Developers should create this input excel file with dummy data for practice purpose. 1003 PDF application can be found at <https://www.fanniemae.com/content/guide_form/1003rev.pdf>. You can also find an online html application at <https://www.secureloandocs.com/agreement.php?id=28788498>. This application will give you information about required fields and other data validations.

Tools / Technologies: Microsoft Excel, ftp

**Scenario 2 – OLTP Data:** Online users can fill and submit the 1003 Mortgage application located at application at <https://www.secureloandocs.com/agreement.php?id=28788498>. When the users submit the application, it is stored in an OLTP database.

Developers should create the OLTP database, populate it with dummy data and get this data periodically by directly connecting to the database or by exporting this data into a flat file.

Tools / Technologies: SQL Server, MS Access, Microsoft Excel, SQL Server Export, ftp

**Scenario 3 – Zillow Data:** RMGreceives 1003 Mortgage applications from Zillow.com market place. Zillow sends data in XMl format to an ftp location.

Developers should create this input XML file with dummy data for practice purpose.

Tools / Technologies: XML, ftp, Textpad editor from <http://www.textpad.com/>

**FTP Information:** (Connect to the FTP location for practice, but the updated project files are located in the CRM *Interview Prep Projects / Mortgage Project (SSIS & SSRS) / Project Files.zip*

FTP Url: [www.novedea.com](http://www.novedea.com)

FTP User: [bitrainee@novedea.com](mailto:bitrainee@novedea.com)

FTP Password: 9729180901

FTP Locations: root/mortgage/excel and root/mortgage/xml

*Note: Feel free to create your own directories if that makes it easy for you*

## R1:SP2: Story 2: Read input data

**Narrative:**

As a Developer

I want to write ETL components that will read various types of 1003 mortgage application input data from various sources

So that I can insert that data into staging tables

**Scenario 1 - Excel Data:** Read excel data from ftp location. Use script task to validate excel to make sure all the required columns are there.

Tools / Technologies: Excel source, Script Component, Microsoft Excel, ftp

**Scenario 2 – OLTP Data:** Read OLTP data from SQL Server or Access database.

Tools / Technologies: OLEDB, SQL Server, MS Access, Excel, SQL Server Export, ftp

**Scenario 3 – Zillow Data:** Read XML data from ftp location.

Tools / Technologies: XML Source, Script Component, XML, ftp

## R1:SP2: Story 3: Insert data into staging tables

**Narrative:**

As a Developer

I want to write ETL components that will insert 1003 mortgage application input data from various sources into staging database

So that I can validate that data for any errors

**Scenario 1 – Staging Database:** Design and create staging tables

Tools / Technologies: SQL Server

**Scenario 2 - Excel Data:** Write excel data into staging database. You may have to transform excel data into format that can be inserted into staging. For example you may have to split some data or combine some data or change the format of some data so that the data can be inserted into staging database.

Tools / Technologies: Script Task, Excel Source, Conversion Components, OLEDB, Views

**Scenario 3 – OLTP Data:** Write OLTP data into staging database. You may have to transform OLTP data into format that can be inserted into staging. For example you may have to split some data or combine some data or change the format of some data so that the data can be inserted into staging database.

Tools / Technologies: Script Task, Conversion Components, OLEDB Source, Views

**Scenario 4 – Zillow Data:** Write XML data into staging database. You may have to transform XML data into format that can be inserted into staging. For example you may have to split some data or combine some data or change the format of some data so that the data can be inserted into staging database.

Tools / Technologies: Script Task, Xml Source, Conversion Components, OLEDB Source, Views

## R1:SP3: Story 4: Validate & Error Handling staging data

**Narrative:**

As a Developer

I want validate the 1003 mortgage application data that is received to make sure that we have all required information in the format that is acceptable

So that I can load the data into OLAP database and data warehouse for reporting.

**Scenario 1 – Read Staging Data and Validate:** Read staging data and validate it for required fields, data types and other types of validations. If there are errors raise errors.

Tools / Technologies: Script Component

Note: Refer appendix for validation requirements.

**Scenario 2 – Do Error Handling:** If there are errors in any application data, please notify loan officer about the error so that they can fix errors.

Tools / Technologies: Script task, Send mail task / SQL command

## R1:SP3: Story 5: Insert data into ODS database

**Narrative:**

As a Developer

I want to write validated 1003 mortgage application input data into ODS database

So that I can create business intelligence reports

**Scenario 1 – ODS Database:** Design and create ODS database.

Tools / Technologies: SQL server

**Scenario 2 – Load ODS Data:** Insert validated data into ODS database.

Tools / Technologies: SP, OLEDB, SQL Server

## R1:SP4: Story 6: Insert data into OLAP / Data warehouse / Dimensional Model

**Narrative:**

As a Developer

I want to write validated 1003 mortgage application input data into data warehouse

So that I can create business intelligence and analytics reports

**Scenario 1 – Dimensional Model:** Design and create SQL Server OLAP / Data warehouse / Dimensional Model (Facts / Dimension).

Dimensions:

1. Dim\_Borrower
2. Dim\_Property
3. Dim\_Loan

Fact:

1. Fact\_Financials

Tools / Technologies: SQL server

**Scenario 2 – Load Dimensions:** Load the following dimensions:

* Fixed dimension and
* Slowly changing dimension.

Tools / Technologies: Stored Procedures, SSIS, OLEDB, SQL Server, Agents

## R2:SP5: Story 7: Generate Reports – Loans to Date

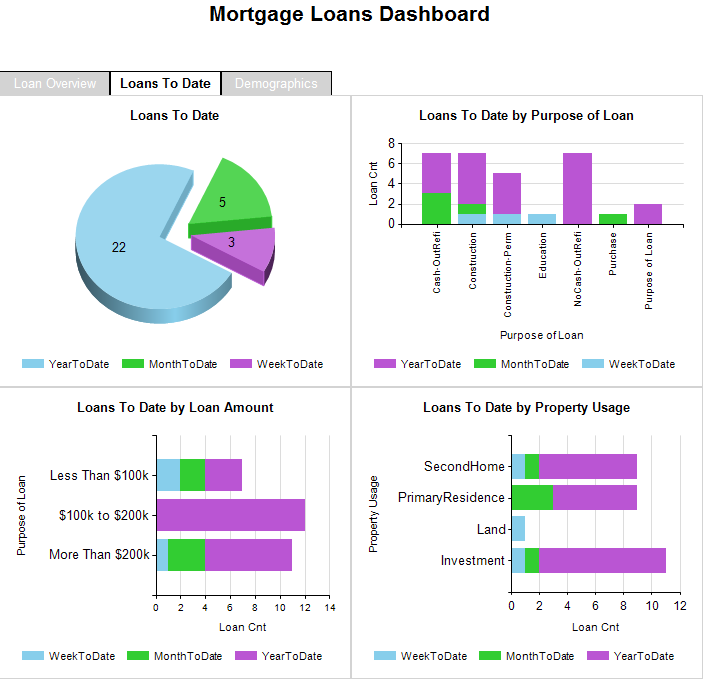
**Narrative:**

As a Loan Department Executive

I would like to have business intelligence analytical reports on a week-to-date, quarter-to-date and year-to-date basis so that I can manage the loan processing business better. I would like to know Loans to date for total loans, purpose of the loan, loan amount breakdown and property usage.

**Scenario 1 – Loans processed to date:** This is the 1st part of the dashboard. Loans processed to Date will consist of 4 graphs

1. Loans to Date
2. Loans to Date by Purpose of Loan
3. Loans to Date by Loan Amount
4. Loans To Date by Property Usage



## R2:SP5: Story 8: Generate Reports – Loans to Date / Demographics

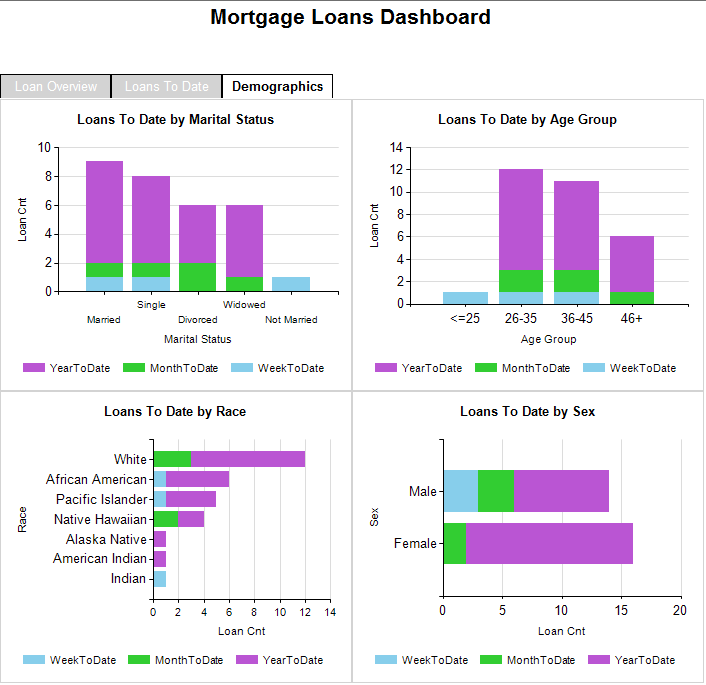
**Narrative:**

As a Loan Department Executive

I would like to have business intelligence analytical reports on a week-to-date, quarter-to-date and year-to-date basis so that I can manage the loan processing business better. I would like to know Loans-to-date for the following demographics: Marital / Status / Race / Sex

**Scenario 2 – Loans processed to date / Demographics:** This is the 2nd part of the dashboard. Loans processed to Date will consist of 4 graphs

1. Loans to Date by Marital Status
2. Loans to Date by Age
3. Loans to Date by Race
4. Loans To Date by Sex



## R2:SP6: Story 9: Generate Reports – Loan Overview

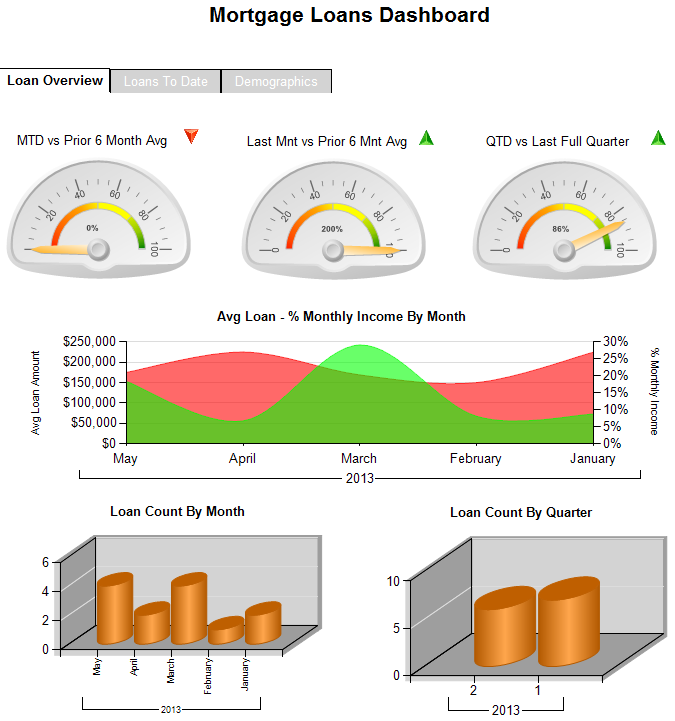
**Narrative:**

As a Loan Department Executive

I would like to have business intelligence analytical reports on the loan overview process so that I can manage the loan processing business better. I would like to know visually appealing dashboard with gauges and different charts.

**Scenario 2 – Loan Overview:** This is the 3rd part of the dashboard. Loans processed to Date will consist of 3 gauges and 3 graphs.

1. Gauges – All gauges will be displayed in %’s. The scale will be from 0 to 100%. The scale will have ranges: Red: 0 to 50% Yellow: 51 to 75% Green 76%+
   1. MTD vs Prior 6 Month Avg
   2. Last Mnt vs Prior 6 Mng Avg
   3. QTD vs Last Full Quarter
2. Avg Loan - % Monthly Income By Month
3. Loan Count By Month
4. Loan Count By Quarter



## R2:SP6: Story 10: Report Functionality

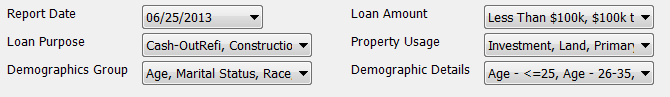
**Narrative:**

As a Loan Department Executive

I would like to have business intelligence analytical reports on the loan overview process so that I can manage the loan processing business better. Preferably, I’d like one interactive dashboard so Story 7/8/9 should be linked together with tabs. I would like the following filters on every report in the dashboard.

*Filters:*

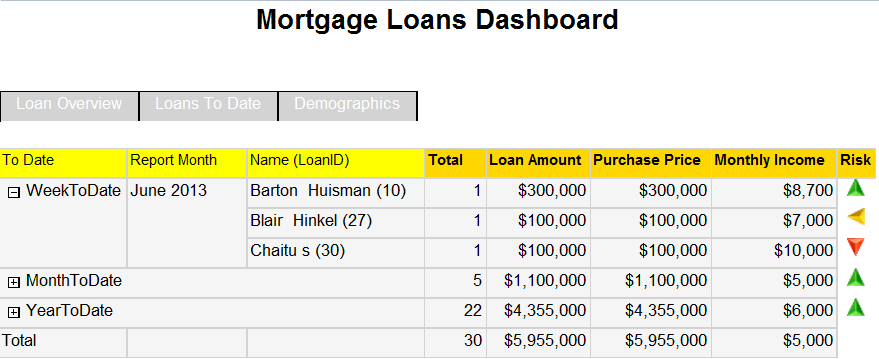
1. **Report Date** – Every loan has a loan date. I want to see a list of all dates and I want the report to default to the most current loan date.
2. **Loan Amount** – Categorize the Loans into 3 groups
   1. Less Than $100k
   2. $100k to $200k
   3. More than $200k
3. **Loan Purpose**
4. **Property Usage**
5. **Demographics Group** – Age / Marital Status / Race / Sex
6. **Demographics Details** – This should display all demographic options. If I select a demographics Group, I would like this list to automatically filter by the attributes in on the groups I selected. For example if I select Marital Status for the Demographics Group filter, this filter should only show the different Marital Status’.
   1. Age – Categorize the Age into the following groups
      1. <= 25
      2. 26-35
      3. 36-45
      4. 46+
   2. Marital Status
   3. Race
   4. Sex



When I select a tab, I want to be able to navigate to the associated report and also pass my same filters that were already selected. Active tables should be white with black font (bold). Inactive tabs should be LightGrey with white font.



When I click on any of the graphs for the *ToDate Report* and the *Demographics Report*, I would like it to drill through to a matrix report so I can see the details of the group I selected. All previously selected filters should also be filtered in the Matrix Report. For example if I click on the Divorced Group in the Marital Status graph in the Demographics report, the matrix report should show all loans from Divorcees along with any other parameter that was included at the time I clicked on the Divorced Group. The matrix will resemble the following picture:



Risk – Calculation based on Monthly Income / Loan Amount.

1. Green Arrow – Sum of Monthly Income / Sum of Loan Amount > 10%
2. Yellow Arrow – Sum of Monthly Income / Sum of Loan Amount between 7 and < 10%
3. Red Arrow - Sum of Monthly Income / Sum of Loan Amount < 7%

***Once completed, the Dashboard should be loaded to your local Report Server where they will be ready for report users and the creation of Automated Subscriptions.***

# Appendix 1: Form Fields Information

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Required** | **Datatype** | **Length** |
| **Personal Info** |  |  |  |
| Borrower FirstName | Y | Varchar |  |
| Borrower LastName | Y | Varchar |  |
| Borrower Email | Y | Varchar |  |
| Social Security number | Y | Varchar | 9 |
| Home Phone | Y | Varchar | 10 |
| Cell Phone | N | Varchar | 10 |
| Marital Status | Y | Varchar |  |
| Date of Birth | Y | Date |  |
| Current Street Address | Y | Varchar |  |
| City | Y | Varchar |  |
| State | Y | Varchar |  |
| Zip | Y | Varchar | 10 |
| Years at this Address | Y | Int |  |
| Base Income(Monthly) | Y | Int |  |
| Bonuses | N | Int |  |
| Commission | N | Int |  |
| Other Income | N | Int |  |
| **Assets:** |  |  |  |
| Checking | Y | Int |  |
| Savings | Y | Int |  |
| Retirement Funds | N | Int |  |
| Mutual Funds | N | Int |  |
| Co-Borrower FirstName | Y (If applicable) | Varchar |  |
| Co-Borrower LastName | Y (If applicable) | Varchar |  |
| Co-Borrower Email | Y (If applicable) | Varchar |  |
| Referral Name | N | Varchar |  |
| Real Estate Agent Name | N | Varchar |  |
| Real Estate Agent Phone | N | Varchar | 10 |
| Real Estate Agent Email | N | Varchar |  |
| Rent or Own | Y | Varchar |  |
| Purpose of Loan | Y | Varchar |  |
| Property Usage | Y | Varchar |  |
| Loan Amount | Y | Int |  |
| Purchase Price | Y | Int |  |
| Number of Units | Y | Int |  |
| Property Address:City | Y | Varchar |  |
| Property State | Y | Varchar |  |
| Property zip | Y | Varchar |  |
| Sex | N | Varchar |  |
| Ethnicity | N | Varchar |  |
| Race | N | Varchar |  |

# Appendix 2: Questions to explore as you do the project

### R1:SP1: Story 1:

* What are various types of sources that you worked with? What is the difference between the sources? What are the issues that you encounter when you work different sources?
* What is the frequency of your source files?
* Have you designed a database? Explain the design of your database.
* What is a parent table and what is a child table?
* What is data dictionary?
* How do you get requirements for ETL projects?
* If there is no business analyst, how do you get the requirement? What is metadata?

### R1:SP2: Story 2:

* How do you load multiple excel files?
* What is ftp task? How do you use ftp task?
* What is control flow?
* Explain for-each look container step-by-step.
* What is a variable?
* What is text qualifier?
* How do you eliminate header rows in flat file?
* What is an expression in SSIS?
* How do dynamically load files in SSIS?
* How do you set the connection string dynamically for excel files or flat files?
* Give me an example for script component / script task?
* What is XML?
* What is XSD?
* What is XSLT?
* How do you load XML files?
* How do you validate XML file if the columns are different?
* What is a web service?
* How do you read web services from SSIS?
* How do you move the files from input to archive?
* When do you use OLEDB? When do you use Native SQL provider?
* What is an Unicode datatype?
* How do you convert Unicode to non-unicode datatype?
* What is bulk insert task?

### R1:SP2: Story 3:

* What are the destinations that you used?
* While loading data, if the column metadata changes what do you do? (refresh metadata)
* What is table lock option? When do you use it?
* Name few staging tables in your application?
* What is referential integrity?
* If you are loading huge amount of data, what are various ways to handle performance bottlenecks?
* How do you performance tuning in SSIS?
* What is check point? When do you use it?
* Can you use check points in dataflow task?
* What is the size of the data that you worked on?
* At what time did you run the utilities packages in your previous project?

### R1:SP3: Story 4:

* How do you validate the data?
* How did you implement error handling in your project?
* Name few data conversions you have done in SSIS.
* What does the derived column do?
* What is conditional split?
* What is multicast, merge join?
* What is the difference between merge in SSIS and SQL?
* How do you parse the null and bad data in SSIS?
* How do you debug data flow?
* What are the asynchronous and synchronous components in SSIS?
* What is the difference between Union ALL and Merge Transformation?
* What is the condition for Merge Transformation?
* How do you sort the data in SSIS?
* Define the row count component. How do you use it?
* How do you send error data to interested parties?
* What will you do if a task fails in SSIS? How do you handle it?
* How do check null values in SSIS? How to replace with other values?
* What is the delay validation?

### R1:SP3: Story 5:

* What is the difference between dimensional model and relational model?
* Define lookup component. Where did you use it?
* What are the cache modes in lookup?
* When do you use a full cache?
* What is the best cache mode for performance? Explain various scenarios.
* If the data is large, do you use lookup?
* What is the alternative for lookup?
* What is merge statement in SQL?
* How do you find out if the data already exists in destination?

### R1:SP4: Story 6:

* What is a fact table?
* What is a dimension table?
* What are the dimensions in your project?
* What are the fact tables in your project?
* How big is your data warehouse?
* What is slowly changing dimension?
* What is Type I, Type II and Type III dimension?
* What are SCD components? How do you load dimensions using SCD component?
* What are the disadvantages of SCD component?
* What are the disadvantages of lookup component?
* Explain merge statement in SQL Server
* If there is large data, how do you load the dimension?
* Which one will load first? Dimension table or fact table? Why?
* What is a surrogate key?
* What is a business key?
* Give some examples of slowly changing dimensions in your project?
* What are the hierarchies in your dimensions?
* What is a cube?
* How do you process the cube?
* What is star schema? What is snowflake schema? Which schema do you prefer?
* What is Kimball methodology?
* What is Inman methodology?
* What is the difference between Kimball and Inman?
* Is star schema normalized?
* What is normalization?

### R2:SP5: Story 7, 8 & R2:SP6: 9,10:

* What are the various reports that you developed?
* What is the difference between tabular reports and matrix reports?
* What is the drilldown report and what is the drill through report?
* Give an example of parameterized report?
* How do you deploy the reports?
* How did you get the requirements for your reports?
* If you do not have a business analyst in your team, how do you get the requirement?
* If your report is running slow, how do you trouble shoot it?
* If the queries are working fine and the report is still running slow, how do you debug the issues and fix them?
* What is the command line utility in the deployment?
* What are the various functions in SSRS?
* What is the function to get todays date in SSRS?
* How do you handle NAN?
* Did you write any custom code in SSRS?
* Have you used any third party DLLs in SSRS?
* What are the various subscriptions in SSRS?
* How do you handle security in SSRS? Give some examples?
* What is a data driven subscription?
* Have you deployed reports in sharepoint? How did you do that?
* What are the new features in 2008 R2 reporting services?
* What are the two installation modes in SSRS?
* How do you deliver reports?
* What is the different between report server and report manager?
* What are the best practices in SSRS in terms of performance tuning?
* What is the complex report that you have developed?
* What is MDX?
* Have you developed any reports from cubes?
* What is power pivot?
* How do you export the reports data into PDF / Excel / Word documents?
* Name few charts you have developed in the past
* What is tablix?
* How do you deploy the reports?
* What is RDL?
* What is report builder?
* Have you done any adhoc reporting?
* Do you have experience working with end users or customers?
* Name the reports other than Tabular, drill down, drill through, parameterized reports?
* Two types of conditional conditions in SSRS?
* What is the expression as getdate() in SSRS?
* What were reports you generated in Saxon?
* Do you have experience in investor reports, business reports?

### Additional Questions:

1. What is transactional data?
2. How do you handle errors?
3. How do you check if the package is running slow and how do you debug the errors?
4. How do you parse the null and the bad data in ssis?
5. Apart from txt files, what other data sources did you deal with? How did you handle the source data other than flat files?
6. What are Kimball and Inman strategies?
7. What do you think are you strong in ssis or ssrs?
8. How do you rate yourself in ssis from 1- 10?
9. When dealing with large data, for matching instead of look up, what alternative do you use?
10. What is your favorite script in script task?
11. Did you ever work on sql 2012 version?
12. What are the configurations did you use in your previous projects?
13. What are the different transformations in the data flow task?
14. What are the reports you generates in SSRS?
15. How did you send mails when errors occurred?
16. Name of the table is Investor, we have investor column , duplicate columns in investor column.
17. With the investor column, select statement that would give only numeric values?
18. What data types would qualify with that?
19. Different between the clustered index and non-clustered index?
20. What is a heap?
21. Difference temp table, table variable?
22. What are the two types of temp tables?
23. Difference between delete and truncate?
24. What are index views?
25. Copy the files from local folder from machine and destination folder?
26. Configure for each loop to achieve the above task?
27. Did you use merge before right?
28. How do you sort the files for best performance?
29. When map is introduced in SSRS?
30. Did you face any problems when developing your packages in any of your project?
31. How did you overcome that scenario?
32. When loading the time dimensions what type of dimension did you use?
33. Did you use accumulating snapshot?
34. What are aggregate tables?
35. What is Partial blocking/no blocking with examples?
36. Lookup Partial cache/full cache/no cache?
37. How to check for null value in ssis/how will you replace it with other value?
38. How to do Transaction in task level in ssis? If there is 8GB memory and your data size is 8GB which cache will you use?
39. How to get records from two sources from in different server?
40. How will you connect those servers?
41. How to get data from two sql server tables using merge join what is the criteria?
42. How do you sort the Records in the table other than using sort component?
43. Have ever worked on DTS/SSIS?
44. If so how will you check whether the output is correct?
45. Can we run multiple dataflow tasks parallelly?
46. What will you do if one tasks failed? How will you handle it?
47. Can you set checkpoint when the tasks are executing parallelly?
48. If you have 10 files how will you load it into Database?
49. What will you do if the package failed in 8th file?
50. There is a tableA every day i'm truncating it and loading it again? the table should not be empty, how will i make sure it?