Plymouth University

School of Computing, Electronics and Mathematics

PRCO304

Final Stage Computing Project

2017/2018

BSc (Hons) Computer Science

Joseph Kellaway

10503639

Cryptocurrency Value Predictor

# Acknowledgements

# Abstract

# Table of Contents

Acknowledgements

Abstract

Table of Contents

1. Background, Objectives, & Deliverables
   1. Concerning Cryptocurrencies
   2. Concerning Blockchain
   3. Introduction
      1. Statement
      2. Motivation
      3. Goals
   4. Overview of Solution
   5. Objectives
   6. Requirements
   7. Initial Scope
   8. Deliverable Parts
      1. Desktop Application
      2. Storage & API
2. Method of Approach
   1. Project Management
   2. Tools
      1. Desktop Application
      2. Storage & API
   3. Version Control
3. Development Stages
   1. Stage 0 – Initial Design
   2. Stage 1 – Data Harvester
   3. Stage 2 – Data Storage
   4. Stage 3 – API & Integration
   5. Stage 4 – GOFAI Algorithms
   6. Stage 5 – GUI
   7. Stage 6 – Benchmarking
   8. Stage 7 – User Trading
4. Evaluation & Testing
   1. MoSCoW
   2. JUnit Testing
   3. User Testing
5. Legal, Social, Ethical, & Professional Issues
   1. Ethics
   2. Security
6. End-Project Report
7. Project Post-Mortem
8. Conclusions
9. References
10. Appendices
    1. Appendix 1 – User Guide
    2. Appendix 2 – Project Management Artefacts
       1. PID
       2. Highlight 1
       3. Highlight 2
       4. Highlight 3
       5. Highlight 4
       6. Highlight 5
       7. Highlight 6
       8. Highlight 7
    3. Appendix 3 – Designs
       1. System Architecture Diagram
       2. UML Diagram
       3. Technology Review
       4. Interface Sketch
       5. Entity Relationship Diagram
       6. Database Normalisation
    4. Appendix 4 – Test Results
    5. Appendix 5 – User Testing Questionnaire

Word count:

# Background, Objectives, & Deliverables

## Concerning Cryptocurrencies

## Concerning Blockchain

## Introduction

### Statement

### Motivation

### Goals

## Overview of Solution

## Objectives

## Requirements

## Initial Scope

## Deliverable Parts

### Desktop Application

### Storage & API

# Method of Approach

## Project Management

## Tools

### Desktop Application

### Storage & API

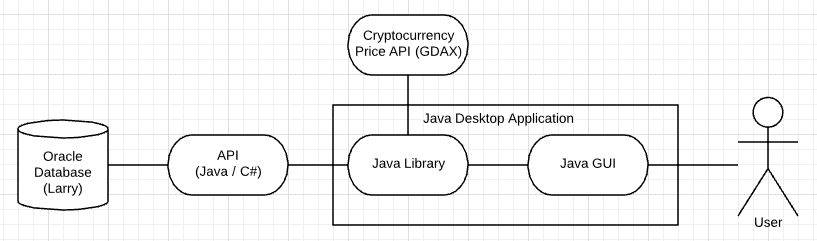
## Version Control

# Development Stages

Development of the project can be segmented into the following stages. This chapter is largely organised in chronological order of development, however, it was logical to separate tasks out modularly and as such some integration took place between stages as necessary.

## Stage 0 – Initial Design

The first prerequisite to starting implementation was understanding how the system was going to work through an entire data cycle. The user would run a Java application that they would interact with in the form of a graphical user interface, this would utilise a library to handle the majority of processing required. The library would collect data from a cryptocurrency exchange’s API endpoint, calculate an average price for the minute, as well as growth from the previous minute and post it to storage that was exposed through an API implementation. When enough recent data is collected, the library would then be responsible for estimating the next value based on recent price changes.



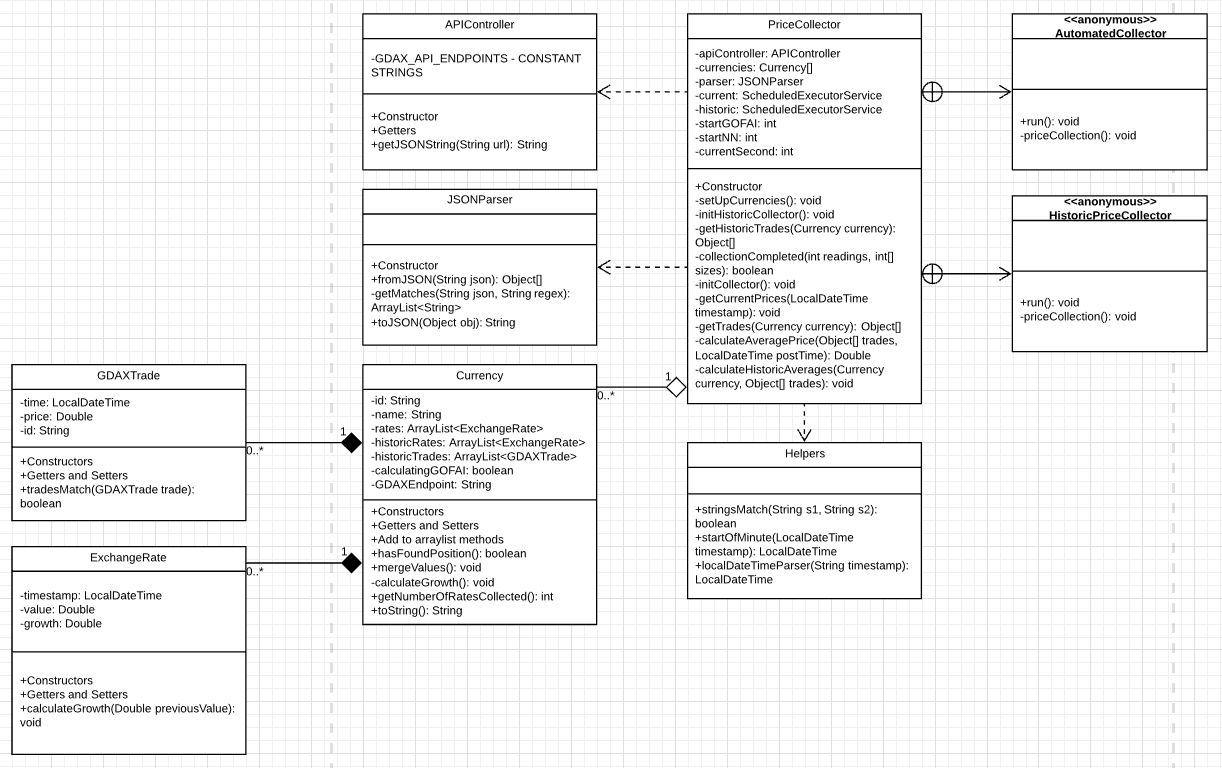
The next requirement would be to determine what data would be stored within the database. This meant briefly sketching the GUI. Due to the fact that the application is intended to be used with minimal monitoring displaying the predictions through the GUI would be illogical. Instead the user would be able to see summarised current statistics for each currency and their trading data for when they are monitoring the program whilst it is trading. There would also need to be a display that allowed them to compare how the trading algorithms are currently performing against other algorithms, against purchasing a single cryptocurrency, and against purchasing no cryptocurrency at all.

GUI Sketch

With these designs in mind, it was possible to start the first stage of development – collecting the data from the chosen cryptocurrency exchange, GDAX.

## Stage 1 – Data Harvester

The initial plan for the data harvester was to have a main PriceCollector class that when initialised would collect Currency information from the database (although, it had to be hard coded at this stage since the database did not yet exist) and then start two collection threads that used the ScheduledExecutorService to periodically submit get requests to GDAX’s API endpoint. The first collection thread would run at the start of each minute to collect a single up-to-date price from the GDAXTrades available. If this thread had not run, then historic data would be collected with a 1 second interval between each batch of get requests to adhere to GDAX’s requests per second limitation.



An APIController would be created to handle communication with GDAX’s API and a JSONParser would parse the JSON strings into GDAXTrade objects. From here the PriceCollector itself would assess which GDAXTrades related to the minute an ExchangeRate was required for, calculate the average price and growth, and add the new rate to the relevant Currency.

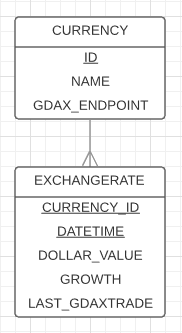
Throughout the project constant monitoring, testing, and improvement of previously implemented systems occurred to ensure that integration didn’t create issues with prior development. This resulted in the PriceCollector’s modification to using only a single ScheduledExecutorService (due to a concurrency issue that occasionally occurred), the APIController being changed to a GDAXAPIController that inherited shared functionality from a new APIController class, the JSONParser was allocated directly to this class and the Helpers class was separated out into a number of helper classes that had a specific functionality (i.e. performing mathematical calculations or modifying a LocalDateTime).

For historic collection batches of prices are collected with each get request (consisting of 100 GDAXTrades) and the PriceCollector averages minute data into ExchangeRates to add to a list of historic prices which is merged when collection is completed. GDAX issues an id number to each trade which is used to traverse pages of trade get requests. The oldest trade id relevant to the ExchangeRate is stored with it to enable the system to have a marker of where it last collected data in the event of the program being turned off, there being a power failure etc.

## Stage 2 – Data Storage

Once data was being collected, brief considerations of GOFAI algorithms were made before it very quickly became obvious that storing collected data would be essential for any algorithm to be meaningfully assessed without wasting vast amounts of time.

Data to be stored was extracted from the proposed GUI interface \*\*(GUI Sketch), then data required to access prices from the GDAX API endpoint was added and normalisation was performed to create a database with the following ERD.



It is shown that each currency may have numerous exchange rates, but only one at a given time. The exchange rate’s compound key consists of the currency’s id as a foreign key and a timestamp and will return the value in United States dollars, growth as a percentage and the last GDAXTrade’s id.

The storage medium used was Oracle as described \*\*earlier and it was created using the create statements (in Appendix 3).

## Stage 3 – API & Integration

## Stage 4 – GOFAI Algorithms

## Stage 5 – GUI

## Stage 6 – Benchmarking

## Stage 7 – User Trading

# Evaluation & Testing

## MoSCoW

## JUnit Testing

## User Testing

# Legal, Social, Ethical, & Professional Issues

## Ethics

## Security

# End-Project Report

# Project Post-Mortem

# Conclusions

# References

# Appendices

## Appendix 1 – User Guide

## Appendix 2 – Project Management Artefacts

### PID

Project Initiation Document

# Contents

|  |  |
| --- | --- |
| Chapter | Page |
| Introduction | 1 |
| Business Case | 1-2 |
| Project Objectives | 2 |
| Initial Scope | 2 |
| Resources and Dependencies | 2-3 |
| Method of Approach | 3 |
| Initial Project Plan | 3 |
| Initial Risk List | 4 |
| Quality Plan | 5 |
| Legal, Social, Ethical and/or Professional Issues | 5 |

# Introduction

Cryptocurrencies such as Bitcoin are a store of monetary value with a volatile price. Purchasing them can result in huge profits or losses of fiat value depending on whether you act at the right time or not. Due to the basic needs of a human being – the need for rest, nourishment, etc. – it is impossible for a single investor to monitor the market all of the time, but with cryptocurrency exchanges never closing, there is a need to do so. It is also possible for a human to apply bias to the decision itself when the need to act arises – favouring currencies that they like the name of, have earned them profitable positions in previous experiences etc. A good night’s sleep coupled with a poor decision can very quickly turn a profitable portfolio into an unsuccessful one.

# Business Case

## Business Need

Currently investors are required to spend a considerable amount of time researching technology behind a particular currency to decide whether to invest or not, potentially missing the optimal opportunity to purchase and therefore missing out on potential profits. Coin prices often spike up and down following news articles, endorsements and criticisms from celebrities, and various other factors including investors simply following the market because they are worried about the value of their portfolio as one currency reduces in value, so they switch to a currency that is holding or likely to increase in value. It is impossible for an individual to follow all potential reasons for price changes simultaneously and calculate their combined influence on the price at any given time, let alone to operate 24/7. It is also possible to ignore such signs due to an affiliation with a particular currency based on past performance.

## Business Objectives

Develop a system that allows for automated trading in a way that:

1. Removes the need for user interaction once the system has been initialised.
2. Allows the user to automatically cash out profits or the initial investment once a target has been reached to prevent loss.
3. Improves trading efficiency by:
   1. Reacting to the market rather than allowing human biases to interfere.
   2. Operating whenever the application is active.
   3. Ensuring no human error causes miscalculations.

# Project Objectives

1. To implement a system that collects data from the GDAX API endpoint and calculates price changes.
2. To implement a process that interprets whether action needs to be taken and acts accordingly.
3. To implement a process that allows the user to protect investment capital.
4. To implement a process that retrains the system as necessary.

# Initial Scope

1. The application collects trading information and calculates the price change percentage against the previous collection, displaying this change to the user.
2. The application is able to make an interpretation of market movements.
   1. The application is able to make a recommendation based upon its interpretation of market movements.
   2. The application is able to act based upon its interpretation of the market movements.
3. The user is able to protect investment:
   1. Option to cash out profits of a set amount when the currency’s value reaches a certain target. (Recurring if constant price rises).
   2. Option to cash out a set amount of the currency once it reaches a certain value. (Single withdrawal to protect investment).
4. The system monitors its prediction accuracy and can be retrained.
   1. Retraining button.
   2. Automated retraining after set amount of time.

# Resources and Dependencies

Space on the Plymouth University Oracle server (Larry) has been requested. Two contingency plans are in place in the event that it isn’t available:

* Space on Xserve is available should I require it because Larry is unavailable.
* I am able to gain access to an old laptop that can be wiped and used, though a free technology will be used (Neo4j is among the candidates).

# Method of Approach

Due to the fact that there are several easily distinguishable tasks within the project, software development will employ an incremental approach with 7 increments:

1. Collect data from source, calculate average prices and store in database.
2. Output recommendation based on recently collected prices.
3. Automated trading based on recently collected prices.
4. Investment protection options.
5. Create recommendation based on more complex system (neural network).
6. System prediction accuracy monitoring including option to retrain neural network.
7. Automated retraining of neural network at set intervals.

Possible technologies are:

* Java
* MySQL
* Oracle

The above list will be compared with alternatives once the project commences.

# Initial Project Plan

|  |  |  |  |
| --- | --- | --- | --- |
| **Stage** | **Start Date** | **Completion Date** | **Details** |
| Initiation | 25/01/2018 | 28/01/2018 | Project initiation document. |
| Technology Evaluation | 29/01/2018 | 29/01/2018 | Review into technologies available; relevance, cost, knowledge etc. |
| Initial High-Level Design | 30/01/2018 | 31/01/2018 | Design documentation; Architecture, database design, user interface sketches. |
| Increment 1 | 01/02/2018 | 08/02/2018 | Increment requirements, design, and implementation; data collection and storage; test results. |
| Increment 2 | 09/02/2018 | 10/02/2018 | Increment requirements, design, and implementation; recommendation; test results. |
| Increment 3 | 11/02/2018 | 18/02/2018 | Increment requirements, design, and implementation; automated trading; test results. |
| Increment 4 | 19/02/2018 | 23/02/2018 | Increment requirements, design, and implementation; investment protection; test results. |
| Increment 5 | 24/02/2018 | 06/03/2018 | Increment requirements, design, and implementation; neural network; test results. |
| Increment 6 | 07/03/2018 | 14/03/2018 | Increment requirements, design, and implementation; retrain neural network; test results. |
| Increment 7 | 15/03/2018 | 23/03/2018 | Increment requirements, design, and implementation; automated retraining; test results. |
| Easter | 24/03/2018 | 15/04/2018 | Catch up / get ahead where possible. |
| System and user acceptance testing | 16/04/2018 | 22/04/2018 | Final testing, ensure final system works, test system with potential end user. |
| Assemble and complete final report | 23/04/2018 | 04/05/2018 | PRCO304 Report. |

## Communication Plan

Review meetings will be held with Marco Palomino (supervisor) each Friday at 15:30. Further communication such as emails may occur as required.

# Initial Risk List

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk** | **Probability** | **Impact** | **Management Strategy** |
| GDAX API discontinuing | Low | High | Source alternative API endpoint. |
| Oracle server goes offline | Low | High | Wait for Oracle/university to resolve, research and consult with supervisor about alternative data storage methods such as a NoSQL database that could be stored locally. |
| Loss of entire project | Low | High | Keep numerous backups. |
| Loss of small portions of project | Medium | Low | Use repository to ensure up to date version accessible. |
| Equipment loss/damage | Medium | Low | Switch from desktop to laptop/spare laptop or use university equipment such as open access. |
| Difficulty with learning requirements mentioned above | Medium | Medium | Apply more time to learning the specific skills required to complete the task. |
| I lack specific knowledge to complete a segment of the project | High | Medium | Accept that all new projects have a learning process attached to them, allow extra time for parts that haven’t been attempted previously. |

# Quality Plan

|  |  |
| --- | --- |
| **Quality Check** | **Strategy** |
| Requirements | It must be ensured that requirements are correct and relevant before implementation.  It must also be ensured that they are complete, achievable, and demonstrable after implementation.  User testing will be used. |
| Design validation | Designs will be checked against requirements.  An entity relationship diagram will be drawn up and database normalisation performed.  Software design principles will be followed. |
| Sub-system verification and validation | To be conducted with testing at the end of each increment. |
| Total system verification and validation | To be conducted at the end of implementation. |

# Legal, Social, Ethical, and/or Professional Issues

The project aims to stick within the limitations of the approved ethics application.

### Highlight 1

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 08/02/2018 |
| **Review of work undertaken**   * A technology review has been carried out, Java has been selected as the primary programming language for the project. * Rough designs for user interface, normalisation, entity relationship diagram and a system architecture diagram have been created. * Development of data collector underway: * Data collection from the GDAX API (RESTful endpoint) commences once per minute and average prices are calculated. Started work to replace with a WebsocketListener. * Data collection for historic data initiates on start-up at once per second, skipping the one second the other collection occurs due to GDAX maximum API requests per second. * Unsuccessful research into alternative exchange APIs that would allow faster collection of historic data.   This week 33.25 hours have been spent working on the project. |
| **Plan of work for the next week**   * Unit testing and Javadoc for implemented data collector. * Implement data storage and use of it pending conversation with supervisor. * Implement and test various price prediction algorithms (GOFAI) to make currency recommendation, select the best one(s) for use. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * No meetings – exchanged emails throughout week |
| **Brief notes from supervisory meeting(s) since last Highlight**   * Marco is happy with my PID. * Contingency plans for data storage were made. |

### Highlight 2

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 15/02/2018 |
| **Review of work undertaken**   * Draw up UML, ERD & architecture diagram using Lucidchart. * Bug fixes for known/discovered issues with data harvester. * JUnit testing for data harvester (including models etc.). Bug fixes uncovered by JUnit testing. * Research into static (GOFAI) stock/cryptocurrency prediction techniques. * Design and implementation of GOFAI prediction algorithm.   This week 36 hours have been spent working on the project. |
| **Plan of work for the next week**   * JUnit test GOFAI algorithm and calculate accuracy. * Integrate selected GOFAI algorithm into (sandbox/spoofed) automated trading, JUnit test. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * No meetings – exchanged emails throughout week |
| **Brief notes from supervisory meeting(s) since last Highlight**   * Upload ERD and system architecture diagrams (completed) |

### Highlight 3

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 22/02/2018 |
| **Review of work undertaken**   * Due to the time required to gather data to test various GOFAI algorithms (and subsequent lost developing time in the event of erroneous code), focus was switched to database integration. * Data normalisation was carried out and Oracle database tables were set up accordingly. * A C# API was implemented to expose the database to the Java application and Java code was designed and is currently being implemented to consume the API endpoints allowing for RESTful processing of the data. Creating the API took longer than anticipated due to conflicts with the Visual Studio version and a reinstallation of both Visual Studio and Oracle Developer Tools was required.   This week 32 hours have been spent working on the project. |
| **Plan of work for the next week**   * Finish Java code to handle RESTful API calls. * Fill database with trade data and test GOFAI algorithms. * Integrate selected GOFAI algorithm into (sandbox/spoofed) automated trading, JUnit test. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * No meetings – exchanged emails throughout week |
| **Brief notes from supervisory meeting(s) since last Highlight**   * None |

### Highlight 4

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 01/03/2018 |
| **Review of work undertaken**   * Modify API GET and PUT calls so that they could be achieved with the current compound key set up. * Expand Java application to attempt to connect to database and act accordingly – if the database connection fails it operates without the use of storage, if the database is empty is begins from scratch and if there is data missing after down time the system begins to fill gap(s) starting with the most recent data. * Use above functionality to fill database with 10,000 prices to enable testing of GOFAI algorithms.   This week 36.75 hours have been spent working on the project. |
| **Plan of work for the next week**   * Test GOFAI algorithms. * Integrate selected GOFAI algorithm into (sandbox/spoofed) automated trading, JUnit test. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * No meetings – exchanged emails throughout week |
| **Brief notes from supervisory meeting(s) since last Highlight**   * None |

### Highlight 5

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 08/03/2018 |
| **Review of work undertaken**   * Added table descriptions to normalisation and uploaded. * Write up MoSCoW method for project and uploaded. * Various bug fixes/logic improvements for price harvester. * Sandbox/spoof trading implementation to test GOFAI algorithms. * Fix more bugs uncovered by spoof trading.   This week 32 hours have been spent working on the project. |
| **Plan of work for the next week**   * Compare GOFAI algorithms and select best performing algorithm over extended period of time. * Develop neural network implementation. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * No meetings – exchanged emails throughout week |
| **Brief notes from supervisory meeting(s) since last Highlight**   * Do user case diagrams or user requirements based on MoSCoW (or any other) method. * Upload normalisation & table description. * Consider using Google Trends or Google Finance Client with the project to monitor the market |

### Highlight 6

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 15/03/2018 |
| **Review of work undertaken**   * Prepare for demo with Marco. * Build and integrate GUI to display functionality. * Rework price collector to single thread due to concurrency issue discovered testing GOFAI algorithms.   This week 33 hours have been spent working on the project. |
| **Plan of work for the next week**   * Implement trading restrictions and issuing trade orders to GDAX endpoint. * OR (pending demo with Marco) * Implement neural network trading predictions. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * 09/03/2018 – 15:30 |
| **Brief notes from supervisory meeting(s) since last Highlight**   * Work towards getting current program in a position to demo it to Marco next Friday. |

### Highlight 7

|  |
| --- |
| **PRCO304: Highlight Report** |
| **Name:** Joseph Kellaway |
| **Date:** 22/03/2018 |
| **Review of work undertaken**   * Connection issues fixed. * Completion of benchmarking process. * Integrate benchmarking process into GUI. * Implement spoof trading – both backend logic and GUI. * Various bug fixes.   This week 32 hours have been spent working on the project. |
| **Plan of work for the next week**   * Implement trading restrictions and issuing trade orders to GDAX endpoint. * OR (pending demo with Marco) * Implement neural network trading predictions. |
| **Date(s) of supervisory meeting(s) since last Highlight**   * 16/03/2018 – 15:30 |
| **Brief notes from supervisory meeting(s) since last Highlight**   * Fix connection issues. Continue working on GOFAI trading mode – do not touch neural network until project cycle complete without it. |

## Appendix 3 – Designs

### System Architecture Diagram

### UML Diagram

### Technology Review

### Interface Sketches

### Entity Relationship Diagram

### Database Normalisation

### Oracle Create Statements

#### PRCO304\_CURRENCY

CREATE TABLE "JOSEPHKELLAWAY"."PRCO304\_CURRENCY"

("CURRENCY\_ID" VARCHAR2(5 BYTE) NOT NULL ENABLE,

"CURRENCY\_NAME" VARCHAR2(30 BYTE) NOT NULL ENABLE,

"GDAX\_ENDPOINT" VARCHAR2(60 BYTE) NOT NULL ENABLE,

CONSTRAINT "PRCO304\_CURRENCY\_PK" PRIMARY KEY ("CURRENCY\_ID")

USING INDEX PCTFREE 10 INITRANS 2 MAXTRANS 255 COMPUTE STATISTICS

STORAGE(INITIAL 65536 NEXT 1048576 MINEXTENTS 1 MAXEXTENTS 2147483645

PCTINCREASE 0 FREELISTS 1 FREELIST GROUPS 1

BUFFER\_POOL DEFAULT FLASH\_CACHE DEFAULT CELL\_FLASH\_CACHE DEFAULT)

TABLESPACE "USERS" ENABLE

) SEGMENT CREATION IMMEDIATE

PCTFREE 10 PCTUSED 40 INITRANS 1 MAXTRANS 255

NOCOMPRESS LOGGING

STORAGE(INITIAL 65536 NEXT 1048576 MINEXTENTS 1 MAXEXTENTS 2147483645

PCTINCREASE 0 FREELISTS 1 FREELIST GROUPS 1

BUFFER\_POOL DEFAULT FLASH\_CACHE DEFAULT CELL\_FLASH\_CACHE DEFAULT)

TABLESPACE "USERS";

#### PRCO304\_EXCHANGERATE

CREATE TABLE "JOSEPHKELLAWAY"."PRCO304\_EXCHANGERATE"

("CURRENCY\_ID" VARCHAR2(5 BYTE) NOT NULL ENABLE,

"DATETIME" TIMESTAMP (0) NOT NULL ENABLE,

"DOLLAR\_VALUE" NUMBER(8,2) NOT NULL ENABLE,

"GROWTH" NUMBER(7,4),

"LAST\_GDAXTRADE" NUMBER(10,0) NOT NULL ENABLE,

CONSTRAINT "PRCO304\_EXCHANGERATE\_PK" PRIMARY KEY ("CURRENCY\_ID", "DATETIME")

USING INDEX PCTFREE 10 INITRANS 2 MAXTRANS 255 COMPUTE STATISTICS

STORAGE(INITIAL 65536 NEXT 1048576 MINEXTENTS 1 MAXEXTENTS 2147483645

PCTINCREASE 0 FREELISTS 1 FREELIST GROUPS 1

BUFFER\_POOL DEFAULT FLASH\_CACHE DEFAULT CELL\_FLASH\_CACHE DEFAULT)

TABLESPACE "USERS" ENABLE,

CONSTRAINT "PRCO304\_EXCHANGERATE\_FK1" FOREIGN KEY ("CURRENCY\_ID")

REFERENCES "JOSEPHKELLAWAY"."PRCO304\_CURRENCY" ("CURRENCY\_ID") ENABLE

) SEGMENT CREATION IMMEDIATE

PCTFREE 10 PCTUSED 40 INITRANS 1 MAXTRANS 255

NOCOMPRESS LOGGING

STORAGE(INITIAL 65536 NEXT 1048576 MINEXTENTS 1 MAXEXTENTS 2147483645

PCTINCREASE 0 FREELISTS 1 FREELIST GROUPS 1

BUFFER\_POOL DEFAULT FLASH\_CACHE DEFAULT CELL\_FLASH\_CACHE DEFAULT)

TABLESPACE "USERS";

## Appendix 4 – Test Results

## Appendix 5 – User Testing Questionnaire