# ABSTRACT

In this section write a bit about what the project is about and the ideal situations that it would be used in.

# Table of Contents

ABSTRACT

Table of Contents

Introduction

Background

System Design

Capter 1

Chapter 2

Chapter 3

Reflexion

References

# Introduction (1page)

## Set the scene of what is currently happening in the word

Athletics athletes throughout New Zealand and around the world have difficulty in competing on the world stage with other High Performing athletes. They all have issues such as backing support for travel or equipment costs. The only effective way of getting support for their discipline is to be performing on a world level stage which unless they develop quickly, they still run into financial troubles when getting to the High Performing level. Sponsors and committies have the ability to support these athletes but they too are stuck in a similar situation. How do they figure out who has the potential to be great and who just wants a free handout.

## Move onto what is out there at the moment

Athletics New Zealand currently have a system in place that targets High Performing athletes within New Zealand and supports them based on their ability using a comparison with High Performing athletes from 1980 - 1995. The program is used only to fund athletes, anything else such as sending athletes to compete at world events goes off an athletes best performance within a certain timeframe and is compared to the times of the same competition prior. Athletes also use the same method of determining where they sit in comparision to other athletes by comparing their best performances done at similar times.

## What this project focuses on. (goals)

This project focuses on collecting, displaying and comparing High Performing athletes with potentially high level athletes. The specific goals include a web scrapping tool for data collection and a graphing program that can display and compare results with user data.

The project takes relevant terms and models used currently in the athletic community to produce visual data that users can relate to and understand.

## What the project consists of. (aim)

The project aims at creating a visual display that both athletes and Sponsors can use to more correctly determine levels of ability based off all performances an athlete has achieved with the added effect of showing future potential of the athlete based off their performances.

## How users will interact

The future intention of the project is to have athletes add performances and compare their abilities not only with high performing athletes but also with athletes of a similar level within their country. Athletes perform on a regular basis so a data base where data can be uploaded and downloaded from will keep athlete programs up-to-date. National level data can also be added in this way giving better comparisons for a wider range of the athletics community.

# Background(Page 2-3)

## BG format.........................

What is happening

Past

Various work

//so doesn’t actually talk about his stuff at all...

## Best of the best

*“Mediocrity acknowledges nothing higher. Talent recognises genius*” – Oscar Wilde.

Countries select athletes to compete for their flag for various reasons. The most common being they are the best in their field for their event. However, what happens when there are 5 or 10 athletes who are all doing the same performances? This is where other factors come in to figure out who will be the best to support and send to world events.

## Factors causing issues

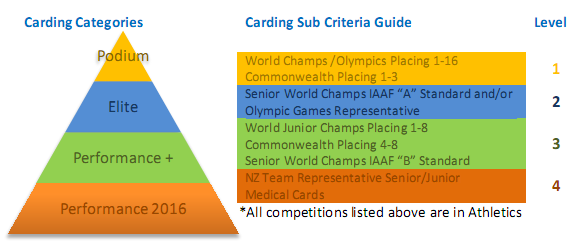
Various factors include but are not limited to; past performances, comparison with world level athletes, mentality, physical well-being, personal history (i.e. Arrests, drug use). Countries can roughly asses from these factors who to support and send based on their results. However, in recent years countries have been attempting to support younger athletes who they believe have the potential to be the next best athletes of the country. This although a great idea does come with disadvantages. Anything could happen to the athlete before they get to the world stage which could see a loss in return from the support offered. They may have hit their peek already or they could decide to give up.

## ANZ

Athletics New Zealand wanted to create a way to support these younger athletes to get them to a high performing level. As a result they created a system to support their athletes as well as a program calculating to make sure they were meeting a certain requirement.

## ANZ System

The system has various rules regarding the selection, keeping the sponsorship and amount of support offered. The basic idea is shown in the pyramid below. There are 4 categories for different levels of athletes. Gold to Green are athletes who have the potential or are in the top 16 in the world for their event. Red is for the younger or slightly lower levelled athletes who are trying to work their way up to the high performing athletes in later years.



## ANZ Program

Their program that they use to calculate what category an athlete should fall works off static data per year and gives a rough estimate off the mean of data from past years performers. It calculates performance levels based off an athletes best performances. The issue here is that by only looking at a select portion of the data available there is plenty of room for error. The data being sampled is at times outlying data meaning that it doesn’t reflect a persons actual ability, only their best ability.

## Steve’s Work

An article from Steve Hollings talks about the simplicity in selecting athletes with potential and the difficulty in coaching and supporting these athletes in the right way to get the most out of them.

<http://www.sportingpulse.com/assoc_page.cgi?client=2-4002-0-0-0&sID=43832&&news_task=DETAIL&articleID=3868164>

## This Project

This project looks at athletes data as a whole and creating a 2nd order polynomial line of best fit through the data. This line is then compared with other high performing athletes to give users an indication of where they are heading in comparison to the worlds best. The benefit of this is it can also give a rough indication of where the athlete is heading with their performances for the future if they stick to the path they’ve been doing....

## What the project will achieve.

The project takes the ideas of Steve Hollings work for comparing High Performing athletes and extends it to be used as a coaching tool as well as producing valuable information around the athlete and their competition or to create a sense of future potential for the athlete to see and work off.

# System Design

## 3 Versions

### How the project is organised

### The project consists of 3 versions of software that have been developed along for a client.

### Talk about all 3 projects and how they have been designed.

## Version 1 Excel VB

The Excel version uses macros and Visual Basic coding to achieve a user friendly environment. It works by opening a form requesting information on the task you want to achieve and then navigating you through the large Excel file to the allocated data set with Graph display.

Macros can be used to setup, format and calculate various aspects of Excel. For this program, the project is already formatted before being given to the user. This is beneficial for two reasons. The first being the processing power on machines. Older or less powerful machines struggle with opening the document as it’s a lot larger than a normal Excel document. The other concern with the size of the document and using Excel is that it loads every sheet at runtime making all sheets available for access. However due to the way users operate the program there is no need for all Excel sheets to be created. When operating the program, you can only ever open one Excel Sheet at a time. As a result, having all sheets loaded causes unnesicary load on the computer, makes the opening and initial usage of the program slow and resulting in people being less inclined to use it.

The ability to store all the data within sheets works to a programmers advantage but holds no real benefit to users. In fact, having data so close to the working area of users increases the probability of them accidentally going into it and destroying the program entirely.

Excel has adequate security however it like most things can be conquered. If a few select people were to be using this software there wouldn’t be an issue with security. However, as this data is used by everyone it’s the equivalent of having all the records stored in a locked room with only a few people having access. If you're sending the file out to people who should only have limited access, that's the equivalent of letting someone into the room and telling them to only look at 1 drawer. They might (and should) do what you said, but there is the chance that they won't. (Luke M. Chandoo.org 2009).

Security for the data is a necessity as without it, the program becomes useless. The client wanted a way of looking after the data collected. You can lock a file but people need to be let in to use the application. You can lock the sheets but the security on this for Microsoft Excel 2007 and earlier versions are very simple to get past.

The layout of the version has distractions reguarding how it looks and how it operates. The issue with creating it in Excel is people are use to the look already and have a hard time focusing on the program when in something else they know how to operate in a different matter.

Features of this version include program directed usability,

## Version 2 Excel C#

### Intro

The interpolation of Excel in C# was done using a browser that navigated to the Excel file directly. How it worked was it would open the requested document and display it in the window. Users then had the ability to use the functions that the Excel document provided while leaving out the Excel Ribbon and titles to the columns and headers of the cells. By itself it proved to do the task that it could do in Excel, however with this version there was no distractions or features being displayed that users weren’t going to use.

### What I didn’t do and why

Another way to read and write data to Excel files using C# is by using OLEDB (Object Linking and Embedding, Database). OLEDB is an API designed by Microsoft to allow accessing data from a variety of sources in a uniform manner. This process allows you to extract and input data into Excel files. The reason I have not used this process in this project is because the client wanted to have the Excel graph not only display the information but also when clicking on it there was a hover effect that could show you the name of the athlete whose data you are viewing. The API can only give you snapshots of updated graphs which although shows the same data, there is no interaction with the graph and user hovers. As a result it was easier to use the excel wrapper to handle the communication with the Excel to C# interpolation than using the API.

OLEDB has a much faster reading a range of cells process than the Excel Wrapper. This is because the API directly works with the data in the Excel file where as the Wrapper is a middle man that gets commands to pass onto the Excel and gets data from Excel to pass onto the C# application.

The reasoning for creating this version was to give it a more professional look. The version before was somewhat thrown together and didn’t really work in with the features already created in Excel. However, after beginning this version a lot more features that were once not available, now could be implemented easily and be used to make the process more effective. Things such as storing multiple athletes for the same event, displaying multiple athletes at the same time and even displaying information about the project could be done with ease.

One of the huge gains from implementing the project this way was the extreme decrease in loading times for the Excel documents. Originally the document was 1.5MB but as the user only ever wants to open up a single page, the new documents that contained only 1 page were 38KB showing a decrease in file size of 97.46%. This means quicker loading times as well as making it easier on older or smaller computers who don’t have much processing power.

MDI proved to be quite challenging. I wanted a way where users could display multiple files at once and either run them together on the display, or have some running in the background on standby while users worked on others. The webbrowser worked well so it seemed like using MDI to display the windows would be very simple to impletemet. This is what happened...

The webbrowser navigates to any Excel document very well however, once navigated to only the last document to be opened has focus. All the rest are only being displayed. If data is entered into one of the windows then the data is stored only in the last document that was opened. An example of this is shown below. The issue was that the webbrowser would only link you to the excel file. As soon as another instance was loaded you were still linked to the file but you could no longer send or receive data from the file any longer. The link to send or receive any new data would be set to the new excel document that was opened.

To solve this issue I needed to make the old versions become the main centre of focus once more by forcing Excel to make the document I was working on to become the main point of focus and put the last one to be displayed to run in the back.

**Excel**

**1**

**2**

**Program**

**3**

**2**

**1**

**3**

Data sent to be stored

## Version 3 C# with Excel

This version was created to work past the Excel boundaries. The main reason for this was to get the flexibility out of the graph. Originally the graph was static and unmoving. If users had data that laid outside the X and Y limits then they were not receiving the full extent of what the program had to offer. The other main reason for the change was the elements that are to be introduced at a later date. Keeping the original excel wrapper would have limited the features and potential displays that the program could produce.

By removing the Excel wrapper a number of improvements became apparent. Firstly, the new graph that is displayed didn’t have any diminishing results. When using the wrapper, there were some small errors that weren’t up to standards/....

The features that are to be worked on a later date are the Funnel error, performance progression of athlete in graph stand outer...

### Closing Excel and other excel issues when working with Interpolation

When closing the program if multiple documents were left open in the program, the wrappers weren’t closing the links they had made to the Excel files. The process was still running even after closing the program. The solution was simple enough. Close the excel file when you close a MDI child. However, MDI forms have a particular way to close child and parent forms. The process works by firing closing all child forms first  
Had I never incorporated this dialogue box into the program, the issue wouldn’t have arose and it would have appeared to have worked exactly as it was doing.

When closing the program if there were children in the MDI forms window the program would ask you if you would like to save the documents and depending on how many were open that many message Boxes would also show. Then it would ask x amount of times again only this time it would close the windows.

Issues:

1. User cant chose to save. Everything is just automatically saved which would make for some sad customers
2. Ignoring means that if the child is closed then the child also ignores the first one which is needed to close it individually.
3. Closed method doesn’t let you cancel out if the user decides that they didn’t want to close the document after all.

The solution was closer to option 2 than I thought. Unfortunately I had gotten too closed minded when working on the problem. It wasn’t until I got a fresh set of eyes to look at the program that the problem was solved. Bill looked at it, looked at my potential solutions as well as the process that the program was working through, saw a pattern and a fix and told me where I was going wrong. Instead of ignoring the first set of closing methods per each form I should have acknowledged the first time they are used and ignore them firing again if they're attempted to be used.

### HCI DESIGN

### The design was that of a gaming menu...

Originally the concept to create the interface was to get a working application that could do everything a user wanted. However, it didn’t occur to me that a user would find how I created it difficult.   
The work done with the client was where the game style intro into the program was created. Users want to be lead through an application the first time, basically having their hand held as they progress through. You see the gaming intro style largely in games, other places that demonstrate this same layout is Microsoft and Apple products such as Excel, Word, Numbers and Presentation.

## GRAPH DEVELOPMENT <-Chapter

* Display of the Excel Version
* Display of the excel version in c# and how you had to change it dues to size alterations. Screen resolution
* Zed Graph and its dynamic functionality

The first version of the graph was done in Excel. The way of displaying the data of the athletes and user data was to use second order polynomials. This was because of...

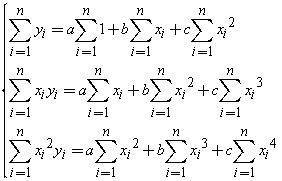
The version that used C# Excel interpolation used a browser to link to the Excel spreadsheet where the graph was displayed. Unfortunatly there was some issues reguarding the window and at the same time there was varying degrees of change in the appearance of graphs on different peoples computers.

Disadvantages of the excel surrounded the amount of data a graph could display comfortably without being too hard on the processing power. The other issue when developing the software was the issue regarding young and / or introductory level athletes. The issue is that the graph was static and adding data outside the graphs displayed axis would not show up becoming useless to these users.

Final resort was to use an open source library zedgraph which has many various features that would solve all issues. The graph was dynamic so you could scroll around to see all the data collected. There is a zoom feature to focus on relevant areas of interest. It like Excel has a way of hovering over a line to get information on that data set. So basically it could tell you what line belonged to who.

The issue with Zedgraph is that it didn’t have the calculations built in to create second order polynomial lines. I ended up coding this myself getting it to work with how zedgraph operates.

### 2nd Order Polynomial



<http://www.efunda.com/math/leastsquares/lstsqr2dcurve.cfm>

Once the 1x3 matrix(A) and 3x3matrix(B) are obtained, a simple calculation is performed to obtain a,b,c matrix (C)

A x B-1= C

Once a,b,c are obtained, you can find the trendline for any amount of data and draw the line using the equation y = ax2+bx+c

### Comparison of two versions

The difference between the way I have calculated how to draw the new trendlines in comparison to how Excel calculates the trendlines are very minute in difference. The graph below shows an example of a plotted graph with random values. The green line shows a trendline drawn off the calculations that Excel has built in. The red line demonstrates the result from the afore mentioned calculation. As you can see there is a slight difference in the dip of the graph as well as where the starting and ending points are drawn. However the difference is so minute that it wont have an affect on the way the program functions at all. As the program works by giving you an estimate of the athletes potential and is used for prediction there is no need for the graph to be 100% precise. The prediction has a standard deviation from the line for error. The other reason is that the future version of the program will be focusing on the error funnel which also wont be affected by the slight difference in the curve at all...

One of the great things about leaving excel to do it in C# was the flexibility it gives you when wanting to compute large sums of data and display this. One of the features that was sort after was a error funnel.

The error funnel works like this...

The error funnel is used for this...

The data collecting will collect many more athletesmeaning potentially more lines occurring on the graph. Excel can only handle so much where as C# is near limitless

Data  
//For this section we are to talk about the threading...

Work around getting the data.

The data used in this project consists of ...

The data was collected by doing this

The data collection now works like this  
-The process of collecting the data was going to use threads and work like this...  
-However now because it’s not reliant on how fast it works but rather if it gets the data or not it does it this way instead.

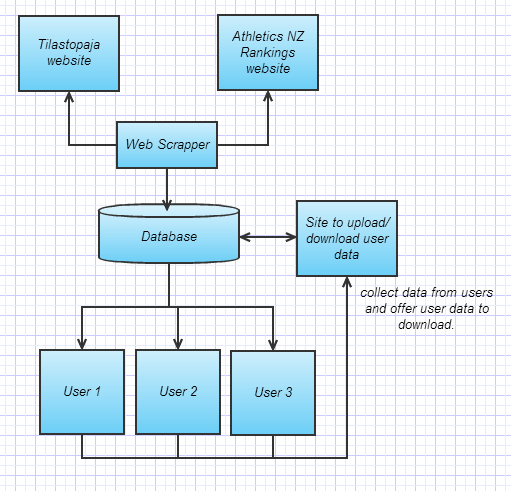
Future methods for data collection include data from Athletics New Zealand HP website as well as data collected from users.

Work around displaying the Data.

The data is still stored in Excel and is done so because...

The database will be done using SQL most likely for web download. Pass on how any of that works but im sure it will be fine.

**Future improvements.**  
Where the project is moving towards...



How this works is, there is a webscrapper that downloads data of High Performing athletes from the Tilastopaja website. It also downloads data from the Athletics NZ rankings website to get National level athletes data. The scrapper stores this information in a database. The main program is distributed to users who get yearly updates of the data. They can then create their own trendlines of athletes performances which can be uploaded to a site for other users to download.

**Why yearly updates...**  
The program only needs yearly or season updates as athletics like many sports is a seasonal activity. During season anyone using this software is going to get little gain from seeing daily or weekly updated trendlines for athletes for two reasons.   
The first is that athletes have already achieved the build up that they are going into the season with. The program doesn’t give you information to work off that will help you each time you put new data in. It works by giving an overview of all your data. Adding one entry is important but generally wont alter the graph alone, the calculations work better with more data.  
The second reason is people who use this software are using it to get a picture of where an athlete stacks up with the rest of the competition at the same age. The data is used as a predictor to calculate whether an athlete is matching up to their fellow competitors and whether they are going to rise or fall in the future based on the results from the past.

**Data privacy**  
As far as data privacy is concerned, users who upload data have the right to remove the data if they so wish. The data that they share and the data collected on High Performing and National level athletes are public source anyway. Any person who performs in a club level or higher event agrees for their results to be displayed to everyone when they perform otherwise it isn’t counted as a legitimate performance.

**Why users want to download other users data**.  
Athletes who are not on a world or national level may want to see what their level is currently doing and compare themselves with their competition. This will generally effect younger athletes of a regional and national level and their coaches. There will be the potential for users to upload fake or misleading data. As this is a possibility it will be up to users to decide who they download from. A rating and comment section could be added to help users download from the best uploaders for better data.

References

(Luke M. Chandoo.org 2009)  
<http://chandoo.org/forums/topic/excel-2007-security-best-practices#post-39583>