# Part A: Data Analysis

## Summary

This report shows an analysis done to determine if differences in certain factors will cause a difference on the mutual fund returns. One-way Analysis of Variance (One-way ANOVA) was used to compare the different categorical groups against the various returns. The results collected are displayed in tables for comparison and an analysis was done to determine which factors could cause any difference in the returns.

## Description of Problem

A data sample of 158 mutual funds was given. Each fund contains 12 variables:

1. Fund: The name of the mutual fund
2. Type: Type of stocks comprising the mutual fund:
   * Small Cap
   * Mid Cap
   * Large Cap
3. Assets: in millions of dollars
4. Fees: Presence of sales charges (No or Yes)
5. Expense Ratio: In percentage of net assets
6. 2001 Return: Twelve-month return in 2001
7. Three-Year Return: Annualized return from 1999-2001
8. Five-Year Return: Annualized return from 1997-2001
9. Turnover: Level of trading activity by the mutual fund
   * Very Low
   * Low
   * Average
   * High
   * Very High
10. Risk: Risk-of-loss factor of the mutual fund
    * Very Low
    * Low
    * Average
    * High
    * Very High
11. Best Quarter: Best quarterly performance from 1997-2001
12. Worst Quarter: Worst quarterly performance from 1997-2001

Based on the variables above, there are three types of returns, namely, the 2001 Return, the Three-Year Return and the Five-Year Return. The project objective is to determine if any of the variables below could cause a difference in the return types:

1. Type
2. Fees

## Description of Data

A summary on the descriptive statistics on the returns are shown in the tables below.

## C:\Users\JianFeng\AppData\Local\Microsoft\Windows\INetCacheContent.Word\means_2001.png

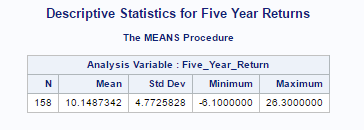
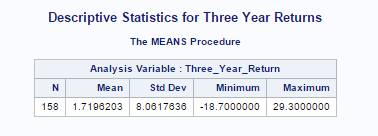


Figure 3.1 Descriptive statistics for the various return types.

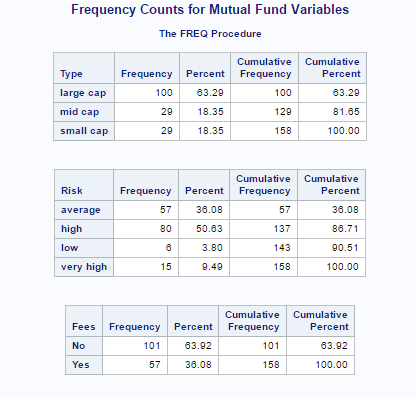


Figure 3.2 Frequency distribution tables for the Type, Risk and Fees variables.

## Discussion of Statistic Analysis Method Used

One-way ANOVA was used for the analysis.

For the variables to be analysed via one-way ANOVA, the numerical variables, namely, the 2001 Returns, Three-Year Returns and Five-Year Returns must fulfil the following conditions explained in the sections below.

### Normality of Variables

The numerical variables must follow a normal distribution for the one-way ANOVA to be valid.

A Shapiro-Wilk test for normality was conducted on the variables. The test was conducted at the 10% significance level. This means that any variable that produces a p-value smaller than 0.1 follows a normal distribution.

The results of the test are shown in the table below.

|  |  |  |
| --- | --- | --- |
| **Test Variable** | **p-value** | **Follow Normal Distribution?** |
| 2001 Return | 5.946e-6 | Yes |
| Three Year Return | 2.254e-5 | Yes |
| Five Year Return | 0.006959 | Yes |

Table 4.1 Shapiro-Wilk Test for Normality on numerical variables.

Based on the test results, all three variables follow a normal distribution.

### Equal Variances

The population variances must be equal across the numerical variables for the group levels.

Bartlett’s Test for Equal Variances was used. The test is conducted at the 5% significance level.

The results of the test are shown in the table below.

|  |  |  |
| --- | --- | --- |
|  | **Fees** | **Type** |
| **2001 Returns** | 0.0003516 | 3.44e-07 |
| **Three Year Returns** | 0.5186 | 0.0001435 |
| **Five Year Returns** | 0.5708 | 0.0001447 |

Table 4.2 Bartlett’s Test for Equal Variances.

### Independence and Randomness of Variables

## Interpretation of Test Results

The one-way ANOVA test results are shown in the table below.

|  |  |  |
| --- | --- | --- |
|  | **Fees** | **Type** |
| **2001 Returns** | F(1, 156) = 3.28, p = 0.0721 | F(2, 155) = 12.57, p = 8.75e-06 |
| **Three Year Returns** | F(1, 156) = 1.179, p = 0.279 | F(2, 155) = 51.99, p = <2e-16 |
| **Five Year Returns** | F(1, 156) = 1.987, p = 0.161 | F(2, 155) = 3.496, p = 0.0327 |

Table 5 One-way ANOVA test for Fees and Type against returns.