**DS-670-Capestone: Big Data & Business Analytics**

**Assignment 2: Data Set Loading and Data Summary**

**January 29th 2017, Madhumita D**

**Data Summary**

Data accumulation and parsing is a recursive process and when we have to analyze financial data then collection of data is an on-going process. The dataset considered is a combination of both profit and loss statement and balance sheet. P&l statement (mostly known as income statement) has indicators which help us in identifying the performance and profitability of a company whereas Balance Sheet has indicators which help us in identifying the financial status of each company. Thus using these indicators from both income statement and balance sheets we try to forecast the stocks with better expected returns.

Every public company issues financial statements quarterly and annually for a fiscal year which includes costs, revenues and expenses. To analyze the profitability and performance of a company it is essential for stock investors to understand the various aspects of these financial statements as they play an important role in whether or not we should invest in a company or stock. The dataset has about 88 indicators for quarterly data on each stock. Each stock or ticker has data on prices, revenue, gross profit, net income, earnings per basic share and other indicators. The expected return value for each stock is calculated for the purpose of this analysis using the quarterly pricing information. We are trying to identify stocks with better expected returns using the indicators in the table below. The final dataset has 20 indicators with quarterly data on 5500 stocks approximately. To this data set we will add the daily closing prices, trading volume and the daily US Treasury yield curve rates for each stock. This data is available on the Data and Charts Center of treasury.gov which coordinates with the U.S. Department of the Treasury.

|  |  |  |  |
| --- | --- | --- | --- |
| **S No** | **Indicator Code** | **Name** | |
| 1 | REVENUE | Revenues | |
| 2 | COR | Cost of Revenue | |
| 3 | GP | Gross Profit | |
| 4 | SGNA | Selling, General and Administrative Expense | |
| 5 | EBIT | Earnings Before Interest & Taxes (EBIT) | |
| 6 | GROSSMARGIN | Gross Margin | |
| 7 | SGNA/Revenue | SGNA as % of Revenues\* | |
| 8 | EBIT/Revenue | EBIT Margin\* | |
| 9 | INTEXP | Interest Expense | |
| 10 | TAXEXP | Income Tax Expense | |
| 11 | NETINC | Net Income | |
| 12 | EBT | Earnings before Tax | |
| **S No** | **Indicator Code** | | **Name** | |
| 13 | EPS | | Earnings per Basic Share | |
| 14 | NETMARGIN | | Profit Margin | |
| 15 | ASSETS | | Total Assets | |
| 16 | ASSETSC | | Current Assets | |
| 17 | LIABILITIESC | | Current Liabilities | |
| 18 | CURRENTRATIO | | Current Ratio | |
| 19 | WORKINGCAPITAL | | Working Capital | |
| 20 | CAPEX | | Capital Expenditure | |

\*Note: SGNA/Revenue and EBIT/Revenue are ratio of indicators by choice.

Each indicator is defined and the relation on it is useful to identify the stocks with better expected returns is described.

1. Revenues: The total amount earned by the company from sale of goods or services after excluding discounts and goods returns is the net revenue and amount earned from sale of goods including the discounts on the goods is the gross revenue. Companies usually state the net revenue in the financial reports. Revenues help us in understanding the growth and performance of the company.
2. Cost of Revenues: It is the total amount of cost incurred by the company in manufacturing or delivering goods or services. Cost of revenue available in the income statement includes the traditional cost of goods sold, direct selling and marketing costs, and yields the lowest margin. If the cost of revenue margin is high then the profitability of a company is high.
3. Gross Profit: Gross profit is the profit a company makes after deducting the costs associated with making and selling its products, or the costs associated with providing its services. Gross profit assesses a company's efficiency at using labour and supplies. It is calculated as revenue minus cost of goods sold.
4. SGNA (Selling, General and Additional Expenses): SG&A is a cost in the income statement which is the sum of all expenses indirectly related to the manufacturing of the product and the daily operations of a company. This is not related to cost of production of the product but the factors which are essential in creation of the product.
5. EBIT (Earnings Before Interest & Taxes) measures the profit a company generates from its operations. It is calculated as revenue minus expenses (excluding the interest and taxes).
6. Gross Margin is a financial measure used to estimate a company's financial health and business model. The amount of money left from revenues after accounting for the cost of goods sold is the gross margin (also known as gross profit margin). It is calculated by dividing the gross profit by revenues. Efficiency of companies is high if we have higher profit margins.
7. SGNA/Revenue: SGNA as a percentage of Revenue is a ratio of the indicators considered.
8. EBIT/Revenue: EBIT Margin is calculated by dividing EBIT by revenues.
9. Interest Expense is the non-operating cost or expense which depicts the interest rates payable for any kind of borrowing funds like loans, bonds, convertible debts. It is the interest accrued over the period of time and not the amount of interest paid.
10. Income Tax Expense is a type of expense which comes under the current liabilities of a company’s balance sheet. It is the taxes which a company needs to pay to the government within a year and is calculated based on the tax law in that country.
11. Net Income: It is the total earnings or profit of a company which is calculated as revenues minus all the costs i.e. cost of business, taxes, interest and depreciation. Net Income is available on the income statement of a company which is used to measure the profitability of a company over a period of time. An earnings per share is calculated using the net income of a company.
12. EBT or Earnings Before Tax shows a company’s financial performance as this is the money retained before deduction the taxes (amount to be paid to the government). It is calculated as revenues minus all the expenses (excluding the taxes).
13. EPS (Earnings per Basic Share) is a profit measure of a company that can be given to one share of its stock.
14. Profit Margin: It is a financial metric which shows how much every dollar of sales of a company is actually kept in the earnings. Profit margin is a profitability ratio which can be calculated by dividing the net income by revenues.
15. Asset is an item with an economic value to the company which can be converted to cash. They can be in any form like securities, cash, inventory and property. They can be categorized into current assets, long term assets, deferred assets and intangible assets.
16. Current assets represent how much capital a company has and are of short term economic value i.e. these assets must be converted to cash with a year. They include accounts receivables, cash and cash equivalents, inventory and prepaid expenses.
17. Current liabilities are available on the balance sheet of a company. They are debts of a company or obligations which legally bind a company to settle a debt within a year.
18. Current Ratio is a measure calculated by dividing the current assets by current liabilities. This ratio ensures whether a company will be able to repay its short-term debts. The financial strength of a company is determined by current ratio.
19. Working Capital: The working capital determines the efficiency of a company as it is calculated as current assets minus current liability. If a company has higher working capital then it is more successful as the money can be invested in expanding or improving operations.
20. Capital Expenditure plays a major role in the growth of a company. The funds used to acquire physical assets such as equipment, property and buildings to improve or maintain the operations of company is considered as capital expenditure.