

# ProfitCents Industry Data

## ABOUT OUR DATA

ProfitCents strives to provide you with the most representative, relevant, and useful data by maintaining our database platform with a uniform, systematic, and statistically valid filtering system. We believe that useful data must be representative, relevant, and timely—our filtering system ensures that the industry averages used throughout our products present you with representative information that accurately reflects the economic conditions of your industry for your financial analysis, whether you segment the data by geographic region, sales range, or data source. First, this paper describes how our cooperative model provides you with timely data. Second, this paper outlines the method behind our filtering process

## BENCHMARK DATA

Our database allows access to real-time benchmark data that we collect directly from our ProfitCents and Sageworks Analyst users through our cooperative data model. This allows us to offer straightforward industry data to our clients who are primarily interested in trending performance metrics for industries and mining industry data through queries. When customers run reports through our ProfitCents and Sageworks Analyst programs, we collect and store that data. We do not divulge identifiable financial information from the specific companies that our users enter into our software. Our database only displays aggregated and anonymous industry-level data. Each day, our customers run over one thousand reports and we incorporate that new data into our industry statistics every morning. Thus, the data that you receive in your report are live. We believe the data from our cooperative model to be the best source of private-company industry data available because they are entered into the system by financial professionals who have an incentive to create the most accurate analyses possible for the business. Despite this incentive, our data are only representative of those private companies that hire financial professionals and, thus, may be limited in scope and/or biased towards more positive results. This potential limitation notwithstanding, the primary purpose of our database is to provide insight into the private sector of the economy—for example, our data are not meant to be predictive of industry or stock market trends.

## THE FIVE FILTERS METHOD

Our five filters method omits inaccurate data from our industry averages. Inaccurate data are neither representative nor relevant—thus, they do not aid your financial analysis. We guard our data for accounting anomalies, input errors, nonsensical testing data, account sample and demo data, and other data inaccuracies that skew our reported averages and common-sized values (see Appendix C for additional information on our common-size adjustments). Before accepting industry data into our database, we screen them using the five filters outlined below:

### Filter 1:

All data are filtered to ensure the validity of the values we collect. We have based these filters on the way that users input financials into our system.

1. No duplicates are included—each set of data must be from a unique business with a unique fiscal period and period length for that business.
2. No data marked as sample data, employee data, or with an unknown data source are allowed.
3. No missing elements are permitted—industry, region, and time period for each financial statement are required.
4. Accepted data must be from audited, reviewed, compiled or tax return-based financial statements.
5. No periods that have been annualized in the program are allowed.
6. Data are not included if they were entered in the “None of the above” industry category or if no industry was indicated.
7. Data that was entered by a customer who chose to opt out of our data collection and aggregation feature are excluded.
8. No data with an invalid date or a future period end date are allowed.

### Filter 2:

We guard against incomplete financial entry—for example, if a customer runs a report wherein all input fields are 0, then we classify this dataset as incomplete and thus inaccurate data. A set of financials with all 0s does not make sense and we omit that report’s values as to not skew our data by lowering our averages with inaccurate 0s.

### Filter 3:

We eliminate metric specific data that violate commonsense accounting rules. We based the following filters on fundamental accounting rules.

1. Net Profit before Taxes cannot be greater than Sales.
2. Adjusted Net Profit before Taxes cannot be greater than Sales.
3. Net Income cannot be greater than Sales.
4. Payroll, Rent, Advertising, Overhead, and Depreciation cannot be less than 0.
5. Cash, Accounts Receivable, and Inventory must be less than or equal to Total Current Assets.
6. Total Current Assets and Net Fixed Assets must be less than or equal to Total Assets.

7. Accounts Payable and Other Current Liabilities must be less than or equal to Total Current Liabilities.
8. Total Current Liabilities must be less than or equal to Total Liabilities.
9. Quick Ratio must be less than or equal to the Current Ratio.
10. Gross Profit cannot be greater than Sales.
11. Interest, Other Operating Income, Other Operating Expenses, Other Income, Other Expenses, Extraordinary Income, and Extraordinary Expenses cannot be less than 0.
12. Other Income / Sales and Other Expenses / Sales cannot be greater than Sales.
13. Taxes Paid / Sales cannot be greater than 1.
14. Other Current Assets must be less than or equal to Total Current Assets.
15. Accumulated Depreciation cannot be greater than Gross Fixed Assets.
16. Other Assets must be less than or equal to Total Assets.
17. Current Portion of Long-Term Debt must be less than or equal to Total Current Liabilities.
18. Long Term Liabilities must be less than or equal to Total Liabilities.
19. Ending Retained Earnings must be less than or equal to Total Equity.

#### Filter 4:

We omit all data that our exception rules generate. Exception rules, which account for, among others, 0s in a ratio's denominator, ensure that we properly display relevant graphs and N/As throughout our products; however, they are not relevant data points per se. We do not believe that ratios generated from exception rules present actionable and representative data for our users—for example, a company with a current ratio of 15500/0 does not present a useful current ratio for that company's industry. We have listed all ratios generated from exception rules in Appendix B.

#### Filter 5:

We omit outliers from our reported averages. Appendix A presents the calculations behind this filter. Suppose, for example, we had dataset where an industry's current ratios were 1.2, 1.1, 1.05, 2.5, 2.9, and 15. In this context, that current ratio of 15 would be an outlier that skews the reported current ratio mean from 1.75 to 3.96—this is a substantial leap for a current ratio. Although that 3.96 is the straight average of this industry's current ratio, that outlier of 15 skews the reported mean so much that the mean would neither be representative nor actionable. Moreover, although 15 is a valid current ratio that one of our customers could have entered, including it in the industry average would not yield accurate financial analysis.

To filter these outliers, we use our two pass standard deviation method to create upper and lower bounds at the five-digit NAICS level—any values that we collect through our cooperative model that fall outside these upper and lower bounds we identify as outliers and we omit from the reported mean. This does not affect how we calculate the industry data, which calculates a new mean from the data within the bounds each time you run a report. Finally, calculating outlier bounds at the five-digit NAICS level ensures that the outlier bounds accurately reflect each industry.

Before we calculate our bounds, we duplicate the five-digit NAICS level dataset and transform that dataset to approximate a normal distribution using a natural logarithmic transformation, which we found to yield the most consistently normal distributions. We then implement an offset function, which combines the logarithmic transformation with a linear transformation, to normalize datasets with 0s and negative values. We then run our first standard deviation script on this normalized dataset to calculate our dataset's lower and upper bounds.

After we calculate bounds and omit the values that fall outside those bounds, we duplicate the dataset and repeat the process again, finally arriving at our industry-specific bounds. This is our two pass method. The second pass picks up outliers that our first pass did not. This is so that we can calculate a more representative average from the dataset that falls within the first standard deviation of any dataset's frequency distribution.

It is important to note that we only omit data—we never delete them from the database. We incorporate that omitted data the next time we calculate our bounds. This ensures that our bounds reflect the entire dataset and not just our already-filtered dataset.

Finally, our team of finance professionals adjusts a small percentage of these bounds, for example, for industries with small sample sizes, to reflect fundamental financial logic. Statistically determined bounds do not always make sense. This is because the standard statistical method +/- two standard deviations creates symmetrical bounds, e.g., a ratio cannot exceed 100 and cannot be less than -100. This does not make business sense for certain ratios, e.g., inventory days. A negative bound for inventory days is not a product of financial logic but rather the product of statistical symmetry—in order to have a negative inventory turnover you must have either negative COGS or negative inventory, both of which do not make business sense.

Thus, we adjust some of the bounds to fit the business scope of our metrics, not just the statistical scope of our metrics. Moreover, if for any reason the two pass standard deviations method produces an unreasonable bound, for example, a gross profit margin upper bound of more than 100%, then we cap those nonsensical bounds at the next most logical result, which for this gross profit margin example would be 100%.

*If you have additional questions on ProfitCents' industry data, please contact us at: 866.4.ABRIGO.*

## APPENDIX A : FILTER 5

For all financial datasets  $\{x_1, x_2 \dots x_i\}$

Define  $Y$  s.t.  $Y = \text{abs}[\min(\{x_1, x_2 \dots x_i\})] + 0.01$

Define new dataset  $\{x'_1, x'_2 \dots x'_i\}$  s.t.  $\forall x_i: x' = \ln(x + Y)$

Define  $\mu_1 = \frac{\sum_{i=1}^n x'_i}{n}$  and define  $\sigma_1 = \sqrt{\frac{\sum_{i=1}^n (x'_i - \mu_1)^2}{n - 1}}$

Define  $UBound_1 = \mu_1 + 2\delta_1$

Define  $LBound_1 = \mu_1 - 2\delta_1$

Define new dataset  $\{x''_1, x''_2 \dots x''_i\}$  s.t.  $\forall x''_i: LBound_1 \leq x'' \leq UBound_1$

Transform dataset  $\{x''_1, x''_2 \dots x''_i\}$  s.t.  $f(x''_i) = e^{x''_i} - Y$

Define  $Y'$  s.t.  $Y' = \text{abs}[\min(\{x''_1, x''_2 \dots x''_i\})] + 0.01$

Define new dataset  $\{x'''_1, x'''_2 \dots x'''_i\}$  s.t.  $\forall x_i: x''' = \ln(x'' + Y')$

Define  $\mu_2 = \frac{\sum_{i=1}^n x'''_i}{n}$  and define  $\sigma_2 = \sqrt{\frac{\sum_{i=1}^n (x'''_i - \mu_2)^2}{n - 1}}$

Define  $UBound_2 = \mu_2 + 2\delta_2$

Define  $LBound_2 = \mu_2 - 2\delta_2$

Define new dataset  $\{x''''_1, x''''_2 \dots x''''_i\}$  s.t.  $\forall x''''_i: LBound_2 \leq x'''' \leq UBound_2$

Transform dataset  $\{x''''_1, x''''_2 \dots x''''_i\}$  s.t.  $f(x''''_i) = e^{x''''_i} - Y'$

Finally, we calculate the average:  $\frac{\sum_{i=1}^n x''''_i}{n}$

## APPENDIX B : ALL EXCEPTION RULES

CurrentRatio	ReturnOnEquity	EBITDAMargin
QuickRatio	ReturnOnAssets	DebtToEBITDA
NetProfitMargin	FixedAssetTurnover	ProfitPerEmployee
AdjNPBTPctChange	GrossProfitMargin	AccountsPayable
SalesPctChange	AdvertisingToSales	AccountsReivable
InventoryDays	PayrollToSales	AccumulatedAmortization
AccountsReivableDays	TotalPayrollToSales	AccumulatedDepreciation
AccountsPayableDays	RentToSales	AdditionalPaidInCapital
InterestCoverage	DebtServiceRatio	AdjNPBT
DebtToEquity		

Advertising	OtherLongTermLiabilities	OverheadToSales
Amortization	OtherOperatingExpenses	CashFlowMargin
Cash	OtherOperatingIncome	OperatingCycle
COGS	OtherStock	OperatingCashFlowMargin
DepreciationToSales	Overhead	LaborCostRatio
CommonStock	OwnComp	SalesPerEmployee
CostOfSales	Payroll	DebtLeverageRatio
CurrentLongTermDebt	PreferredStock	AssetComposition
Depreciation	Rent	ReturnOnLabor
DirectLaborToSales	RetainedEarnings	ProfitPerEmployee
DirectMaterialsToSales	SeniorDebt	CashFlowMarginPctChange
EBITDAMargin	ShortTermDebt	Debt to Tangible Net Worth Ratio
GrossFixedAssets	SubordinatedDebt	Senior Debt to Cash Flow
GrossIntangibleAssets	Taxes	Debt to Cash Flow
GrossProfit	TotalCurrentAssets	Debt to Capitalization
Interest	TotalCurrentLiabilities	Operating Profit Margin
Inventory	TotalEquity TotalLiabilities	Sales Growth Rate
LongTermLiabilities	WorkingCapital	Net Income Growth Rate
NetFixedAssets	InventoryDaysPctChange	Notes Payable / Current Portion of Long Term Debt
NetIncome	AccountsReceivableDaysPctChange	Notes Payable / Senior Debt
NetIntangibleAssets	AccountsPayableDaysPctChange	Notes Payable / Subordinated Debt
NPBT	InterestCoveragePctChange	Dividends Paid / Withdrawals
OperatingProfit	DebtToEquityPctChange	WorkingCapitaltoAssets
OtherAssets	ReturnOnEquityPctChange	G&APayrolltoSales
OtherCurrentAssets	ReturnOnAssetsPctChange	
OtherCurrentLiabilities	FixedAssetTurnoverPctChange	
OtherExpenses		
OtherIncome		

## APPENDIX C : COMMON-SIZE ADJUSTMENTS

Our data comes in from a variety of Financial Statement sources, with different levels of detail in regards to the number and granularity of sub-accounts. In order to remove bias and provide consistency between common-size metrics, we follow an adjustment procedure for data sets of more than 25 statements, which we believe is the threshold for statistical significance. Below are the rules that we ensure to always hold:

### ***Income Statement:***

1. Sales (100%) - Cost of Sales = Gross Profit
2. \* Gross Profit = Depreciation + Overhead - Other Operating Income + Other Operating Expenses + Interest - Other Income + Other Expenses + Taxes + Net Income + Amortization
3. Operating Profit = Gross Profit - Depreciation - Overhead + Other Operating Income - Other Operating Expenses - Amortization
4. Net Profit before Taxes = Operating Profit - Interest + Other Income - Other Expenses
5. Adjusted Net Profit before Taxes = Net Profit before Taxes + Owners Compensation
6. EBITDA = Net Income + Interest + Depreciation + Amortization + Taxes
7. Net Income = Net Profit before Taxes – Taxes

### ***Balance Sheet:***

1. \* Total Assets (100%) = Total Liabilities + Total Equity
2. \* Net Fixed Assets = Gross Fixed Assets - Accumulated Depreciation
3. \* Total Assets (100%) = Total Current Assets + Net Fixed Assets + Other Assets
4. \* Total Liabilities = Total Current Liabilities + Total Long Term Liabilities
5. \*\* Total Current Assets = Cash + Accounts Receivable + Inventory + Other Current Assets
6. \*\* Total Current Liabilities = Accounts Payable + Current Portion of Long Term Debt + Other Current Liabilities + Short Term Debt
7. \*\* Total Equity = Retained Earnings + Common Stock + Preferred Stock + Additional Paid-in-Capital + Other Stock

The sub-accounts under the main accounts indicated with asterisks (\*) are summed up to arrive at a certain value. If this total value does not equal the value of the main account, each of the sub-accounts is multiplied by the ratio of the value of the main account divided by the total summed value. The sub-accounts under the main accounts indicated with two asterisks (\*\*) are adjusted by the same ratio that was used to adjust the main accounts in prior steps.

For example, let's look at the equation Net Fixed Assets = Gross Fixed Assets - Accumulated Depreciation.

Let's say that:

Net Fixed Assets = \$84,000

Gross Fixed Assets = \$100,000

Accumulated Depreciation = \$20,000

Then,

$$\text{Gross Fixed Assets} * \frac{\text{Net Fixed Assets}}{\text{Gross Fixed Assets} - \text{Accumulated Depreciation}} = \$ 100,000 * \frac{\$ 84,000}{\$ 100,000 - \$ 20,000} = \$ 105,000$$

$$\text{Accumulated Depreciation} * \frac{\text{Net Fixed Assets}}{\text{Gross Fixed Assets} - \text{Accumulated Depreciation}} = \$ 20,000 * \frac{\$ 84,000}{\$ 100,000 - \$ 20,000} = \$ 21,000$$

$$\text{Net Fixed Assets} = \$ 105,000 - \$ 21,000 = \$ 84,000$$