

# OSHA Data Initiative's (ODI) Meaning and Findings

Scott Arias  
Eastern Kentucky University  
[Scott.Arias@eku.edu](mailto:Scott.Arias@eku.edu)

## Abstract

During a research study conducted in 2013 several research findings became apparent about safety statistical data provided by OSHA. These research findings, using data from the 2008 OSHA Data Initiative (ODI), included analysis of mishaps by organizational type, ownership, annual revenue and numerous other organizational wide variables that were studied. These findings serve as the basis for conclusions that will be presented, primarily focused on the statistical data meanings and the findings that have relevance to the safety within the industrialized world.

## Introduction

In 2013 a quantitative research study was performed on the relationship between safety performance and the use of the Engineering Manual 385 [1]. The premise of this research study was based on the lack of documented evidence linking safety planning to fewer and less severe mishaps [2]. This research examined three common safety metrics, which included the total number of recordable cases rate (TCR), the days away, restricted, and transferred (DART) rate, and the days away from work (DAFWII) rate. These three metrics help quantify the number and severity of mishaps for each industry within the United States. To analyze this data a structural equation modeling technique (multiple regression) was used.

This 2013 study used construction contractors from the OSHA Data Initiative (ODI) for the 2008 calendar year. The construction contractor mishap data gathered from ODI contained information on the three dependent variables (TCR, DART, and DAFWII rates) [3]. This data was merged with the 2008 federal spending database, which contained numerous explanatory variables. The data was broken down into 5 separable influences, which included the safety protocol utilized, the project location by OSHA region, the size of the contractor by the number of employees, the type of business ownership and the project solicitation and pricing procedures.

## Data

The OSHA Data Initiative (ODI) obtains data annually from randomly selected companies within targeted industries, due to their high frequency of mishaps [4]. The data that this research utilized provided the basis for assessing the effectiveness of the EM 385, within the 2008 calendar year. This research used three dependent variables from ODI which included the total case rate of mishaps (TCR), the days away, restricted, and transferred (DART) and the days away from work (DAFWII). For the use in this research the TCR reflected the number of mishaps, while DART and DAFWII reflected the severity of those mishaps [5]. In

the ODI database the TCR, DART and DAFWII were available by company name and address, with the data normalized to 100 worker. While ODI does say the data is not statistically randomly sampled, it uses the output of the data to provide guidance to industries [6].

In analysis of the data NAICS 236220, 236210, 236116, 236115 was collected from the 2008 OSHA Data Initiative [7]. This data was merged by contractor with the result of the database contained all project for contractors performing work within fiscal year 2008. . From this dataset, the projects were extracted by funding source for use as a determining factor on whether the EM385 was used or not. Using dichotomous variable the funding sources were coded. Due to a lack of independent variables, the federal spending database was utilized by merging the descriptive data with the dependent variables. This was a manual process due to fact these databases are prepared independently. In total there were over 40,000 observations that data analysis was performed on.

The proposed model included the following dependent and independent variables within the research equation:

$$Y_i = \beta_0 + \beta_j (EM385_j) + \alpha_k (REGION_k) + \gamma_l (REV_l) + \delta_m (EMPL_m) + \zeta_n (SIZE_n) + \eta_p (PTYPE_p) + \theta_q (CTYPE_q) + \lambda_r (SOLIC_r) + \xi_s (DIS_s) + \pi_t (SET_t) + \varepsilon_0$$

Listed below are the dependent and independent variables definitions:

Yi=1	= TCR
Yi=2	= DART
Yi=3	= DAFWII
$\beta_0$	= Intercept
EM385 j=1	= EM385 - Major Funding Agency Category
EM385 j=2	= EM385 - Major Agency Category of Contract
REGION k=1	= Place of Performance OSHA Regions 1, 2, 3 & 5 - North East
REGION k=2	= Place of Performance OSHA Regions 7, 8, & 10 - North West
REGION k=3	= Place of Performance OSHA Region 4 - South East
REGION k=4	= Place of Performance OSHA Region 6 & 9 - South West
REV l=1	= Contractor Size by Revenue (\$)
EMPL m=1	= Contractor Size by Employees (#)
SIZE n=1	= Project Size (\$)

PTYPE p=1	= Project Type (Residential)
PTYPE p=2	= Project Type (Commercial)
CTYPE q=1	= Contract Type (Unit Price)
CTYPE q=2	= Contract Type (Lump Sum)
SOLIC r=1	= Solicitation Procedures (Negotiated)
SOLIC r=2	= Solicitation Procedures (Competitive Bid)
DIS s=1	= Disadvantaged Business
DIS s=2	= Non-Disadvantaged Business
SET t=1	= Set-Aside Used
SET t=2	= Set-Aside Not Used
$\epsilon_0$	= Error

The dependent variables included:

TCR - Total number of recordable mishaps as indicated on the OSHA 300A Form.

DART - Days away, restricted, and transferred as indicated on the OSHA 300A Form. The more severe the mishap, the greater the DART.

DAFWII - Days away from work as indicated on the OSHA 300A Form. The more severe the mishap, the greater the DAFWII.

\* These common metrics are normalized to 100 full-time employees.

The independent variables included:

EM385 - Major Funding Agency Category: Federal Department that funds & manages construction work (i.e. Department of the Army) [8].

EM385 - Major Agency Category of Contract: Federal Agency that obtains the construction funding from congress (i.e. Department of Defense) [8].

OSHA Region: OSHA region by project location (10 Regions).

Contractor Size by Revenue: Contractor size by the annual company gross revenue.

Contractor Size by Employees: Contractor size by the number of employees employed.

Project Type: Construction sector of project (Residential or Commercial).

Contract Type: Method of procurement by contract type (Lump Sum or Unit Price).

Solicitation Procedures: The number of contractors that bid a construction project (i.e. Negotiated or Competitive Bid)

Disadvantaged Business: The disadvantaged business status of contractor performing the construction work [9].

Set-Aside: The type of limited procurement method used by the federal government to obtain a construction contract (i.e. Minority-Owned, Veteran-Owned, etc...) [9].

With the variables identified a correlation matrix was completed to assess multicollinearity issues. With a Pearson correlation of .60 or greater, there were two issues that were identified. The first was all dependent variables were correlated. Secondly disadvantaged businesses and set-aside were also highly correlated. Both of these correlations were expected and indicated that our data collection process was performed accurately.

Listed below are the descriptive statistics and assessment used in this research:

**Table 1: EM385 Proxy**

		EM385 Proxy			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Do not use EM385	16	8.4	8.4	8.4
	1 – Use the EM385	174	91.6	91.6	100.0
	Total	190	100.0	100.0	

After reviewing the correlation matrix, the EM385 Proxy was analyzed. This is an explanatory variable that asserts the use or non-use of the EM385. The result was that 91.6% used the EM385, while 8.4% did not use the EM385.

**Table 2: Contractor # of Employees**

Contractor # of Employees					
	Number of Employees Per Hundred	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	30	15.8	15.8	15.8
	.01 - .10	109	58.3	58.3	74.2
	1.01 – 2.00	13	6.7	6.7	78.9
	2.01 – 3.00	6	3	3	81.6
	3.01 – 4.00	4	2	2	83.7
	4.01 – 5.00	1	.5	.5	82.1
	5.01 – 10.00	12	6.1	6.1	88.9
	10.01 – 20.00	6	3	3	94.2
	20.01 – 50.00	3	1.6	1.6	95.8
	50.01 -100.00	4	2.1	2.1	97.9
	100.01 - 500.00	2	1	1	98.9
	501.00 – 1000.00	1	.5	.5	99.5
	1000.00 – 2013.66	1	.5	.5	100.0
	Total	190	100.0	100.0	

After reviewing the correlation matrix, the Contractor Size by Employees was analyzed. The result was that the companies that had from 1 to 100 made up 58%, while companies that had 100-200 made up 6.7% and companies with 500-1000, made up 6.1%.

**Table 3: Solicitation Procedures**

Solicitation Procedures					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Negotiated Bidding	47	24.7	24.7	24.7
	1 – Competitive Bidding	143	75.3	75.3	100.0
	Total	190	100.0	100.0	

After reviewing the correlation matrix, the solicitation procedures were analyzed. The result was that most companies were awarded projects based on competitive bid (76%), with far fewer were awarded by negotiated bid (25%).

**Table 4: OSHA Regions**

OSHA Region 1					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 1	177	93.2	96.2	96.2
	1 – Region 1	7	3.7	3.8	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 2**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 2	174	91.6	94.6	94.6
	1 – Region 2	10	5.3	5.4	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 3**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 3	162	85.3	88.0	88.0
	1 – Region 3	22	11.6	12.0	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 4**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 4	139	73.2	75.5	75.5
	1 – Region 4	45	23.7	24.5	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 5**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 5	160	84.2	87.0	87.0
	1 – Region 5	24	12.6	13.0	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 6**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 6	145	76.3	78.8	78.8
	1 – Region 6	39	20.5	21.2	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

**OSHA Region 7**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 7	172	90.5	93.5	93.5
	1 – Region 7	12	6.3	6.5	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

OSHA Region 8					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 8	172	90.5	93.5	93.5
	1 – Region 8	12	6.3	6.5	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

OSHA Region 9					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 9	173	91.1	94.0	94.0
	1 – Region 9	11	5.8	6.0	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

OSHA Region 10					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not OSHA Region 10	182	95.8	98.9	98.9
	1 – Region 10	2	1.1	1.1	100.0
	Total	184	96.8	100.0	
Missing	System	6	3.2		
Total		190	100.0		

After reviewing the correlation matrix, the location of the project was analyzed. The results were that OSHA Region 4 & 6 were the largest. This is due to the military installations in southeast and southwest sections of the United States. On the other hand the smallest region were in OSHA Region 10, which is 1.1% and is remote locations with few installations.

**Table 5: DBE**

Disadvantaged Business (DBE)					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 – Not DBE	113	59.5	59.5	59.5
	1 - DBE	77	40.5	40.5	100.0
	Total	190	100.0	100.0	

After reviewing the correlation matrix, disadvantaged businesses were analyzed. The results were that Disadvantaged Business Entities (DBE) were 40% of the population, while 60% were not DBEs.

After running hundreds of models, it became apparent that both significant variables and ones that were essential to the research should be retained. From that the final model became apparent which excluded only one variable. This variable was Set-Aside, since Disadvantage Businesses and Set-aside were highly correlated.

With the final model a regression analysis was completed and here are the results. First R2 was assessed, which showed us our model has the ability to predict 22%, 13% and 12%, respectively. These were statistically very low which was a concern, but given the fact that

dummy variables were extensively used and the coefficient results showed issues with the data the analysis continued. These percentages show the variation in the dependent variables can be attributed to the explanatory variables used in this model.

**Table 6: Model Summaries**

**TCR Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.472 <sup>a</sup>	.223	.163	5.30296

**DART Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.363 <sup>a</sup>	.132	.065	3.43742

**DAFWII Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.355 <sup>a</sup>	.126	.059	2.23218

Looking at the coefficient tables, there were 3 items that were focused on. The first is the sign of the coefficient (Showing us whether the DV (Number of mishaps) increase or decreases mishap rates). Second we are looking at the coefficient value for the impact the coefficient has on the dependent variables. Finally we look at the significance, telling us whether the other two should be considered. So when we look at the coefficients in the three models shown below it can be seen that disadvantaged businesses and a few OSHA regions show significant in sign, impact and significance.

**Table 7: TCR Coefficients**

TCR Coefficients<sup>a</sup>

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	2.044	1.796		1.138	.257
EM385 Proxy	1.203	1.491	.057	.807	.421
Contractor # of Employees	-.001	.003	-.031	-.446	.656
OSHA Region 1	6.232	2.211	.206	2.818	.005
OSHA Region 2	-.691	1.889	-.027	-.366	.715
OSHA Region 3	-1.417	1.416	-.080	-1.001	.318
OSHA Region 4	-1.309	1.167	-.097	-1.122	.263
OSHA Region 5	1.907	1.385	.111	1.377	.170
OSHA Region 7	1.141	1.778	.049	.642	.522
OSHA Region 8	6.620	1.753	.283	3.777	.000
OSHA Region 9	-.911	1.819	-.037	-.501	.617
OSHA Region 10	9.707	3.857	.174	2.517	.013
Solicitation Procedures	.511	.937	.038	.545	.587
Disadvantaged Business	1.962	.829	.166	2.367	.019

**Table 8: DART Coefficients**

DART Coefficients<sup>a</sup>

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	1.093	1.164		.939	.349
EM385 Proxy	.761	.967	.059	.787	.432
Contractor # of Employees	.000	.002	-.009	-.118	.906
OSHA Region 1	2.675	1.433	.144	1.866	.064
OSHA Region 2	.048	1.224	.003	.039	.969
OSHA Region 3	-1.170	.918	-.107	-1.275	.204
OSHA Region 4	-1.166	.756	-.141	-1.542	.125
OSHA Region 5	.576	.898	.055	.642	.522
OSHA Region 7	.067	1.153	.005	.058	.954
OSHA Region 8	2.380	1.136	.166	2.095	.038
OSHA Region 9	-1.096	1.179	-.073	-.929	.354
OSHA Region 10	-2.679	2.500	-.078	-1.071	.286
Solicitation Procedures	.448	.607	.054	.738	.461
Disadvantaged Business	1.203	.537	.166	2.238	.027

**Table 9: DAFWII Coefficients**

DAFWII Coefficients <sup>a</sup>					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	.016	.756		.021	.983
EM385 Proxy	.648	.628	.077	1.033	.303
Contractor # of Employees	.000	.001	-.027	-.369	.713
OSHA Region 1	2.249	.931	.187	2.417	.017
OSHA Region 2	1.431	.795	.141	1.800	.074
OSHA Region 3	-.184	.596	-.026	-.309	.758
OSHA Region 4	-.302	.491	-.057	-.614	.540
OSHA Region 5	.688	.583	.101	1.180	.240
OSHA Region 7	.334	.748	.036	.446	.656
OSHA Region 8	.563	.738	.061	.764	.446
OSHA Region 9	-.663	.766	-.068	-.866	.388
OSHA Region 10	-1.301	1.624	-.059	-.801	.424
Solicitation Procedures	.251	.394	.047	.637	.525
Disadvantaged Business	1.023	.349	.219	2.932	.004

After a methodical and through review of data, there was no conclusive quantitative evidence that the EM385 use does or does not result in fewer or less severe mishaps. However, we have learned much in the examination and combination of over 40,000 records of data from OSHA and Federal Spending Database.

**Results and Recommendations**

From this 2008 research study several issues became apparent. Rather than discovering the link between safety planning and a reduced number and severity of mishaps, the data analysis uncovered issues with the data itself. The following recommendations for policy changes could help increase the reliability of the data collected, since currently the data is not statistically useful:

**Policy Change #1**

The data collected and provided by ODI lacks key descriptive data in order to meaningful conduct quantitative research. This hinders researchers conducting quantitative research and uncovering conclusions that could help reduce the number and severity of mishaps. The laborious process of manually linking mishap data with federal spending database hindered the ability to provide quantitation data analysis. Because of this it is recommended that OSHA change its data collection to encompass descriptive data within the OSHA 300A form that is used to collect data.

**Policy Change #2**

ODI should collect additional contractor mishap data. With the abilities of modern day technology this could be achievable and can reduce the difficulty in submitting their OSHA 300A form. More data could assist in more comprehensive quantitative research results.

### Policy Change #3

OSHA should audit the data that is collected. The quality of the data analyzed is the big discovery uncovered by this research. Both ODI and the federal spending database both had many pieces of data that were omitted or obviously erroneous. The quality and accuracy of the data was a major concern of this research with the research model having a strong theoretical basis. None of the regressions performed as part of this research provided for any definitive and significant quantitative conclusions, and this is directly attributed to the lack of quality of the data provided by ODI and the federal spending database.

### Policy Change #5

Workers compensation claims (EMR data) should be linked to the data submitted to ODI. This would require an OSHA policy change to ensure the OSHA 300A matches the claims submitted to workers compensation insurers. These research results call into question the quality of the data collected. The quality of the data can only be increased through a process of audits. The result of this audit will be useful quantitative research results and management outcomes.

### Summary

In summary, this research revealed that the 2008 data used from two separate databases were of questionable quality. This suggested that any future research using this data will be ineffective without the policy changes cited previously. With the large amount of time, money and resources used by the federal government to gather this information, either the data quality needs to be improved or the data collection should be ceased. It is recommended future research not utilize the government resources used in this research until some of the policy changes have had a chance to take effect. All of the results of this research point to one conclusion and that is the data is not of the quality to provide any significant contribution to research. The model initially proposed and is founded in numerous research studies could provide significant results if changes in data collection and auditing policies were corrected.

### References

- [1] U.S. Army Corps of Engineers (2008). *Safety and health requirements manual (EM 385-1-1)*. Retrieved from [www.usace.army.mil/](http://www.usace.army.mil/)
- [2] Hare, B., Cameron, I., & Roy Duff, A. (2006). Exploring the integration of health and safety with pre-construction planning. *Engineering, Construction and Architectural Management*, 13(5), 438-450. doi: <http://dx.doi.org/10.1108/09699980610690729>
- [3] Weil, D. (2001). Assessing OSHA performance: New evidence from the construction industry. *Journal of Policy Analysis and Management*, 20(4), 651-674. Retrieved from <http://proquest.umi.com>
- [4] IN.gov, n.d. Retrieved from <http://www.in.gov/dol/2340.htm>
- [5] Occupational Safety and Health, n.d. Retrieved from <https://www.osha.gov/osha40/timeline.html>
- [6] Bradbury, J. C. (2006). Regulatory federalism and workplace safety: Evidence from OSHA enforcement, 1981-1995. *Journal of Regulatory Economics*, 29(2), 211-224. doi:<http://dx.doi.org/10.1007/s11149-006-6036-1>

- [7] U.S. Department of Labor, Bureau of Labor and Statistics. (2008). *Injury cases - rates, numbers, and percent relative standard errors - by industry – 2009*. Washington, D.C.:Author. Retrieved from <http://www.bls.gov/>
- [8] DeCoopman, Daniel L. (2011) “The Silver Bullet: U.S. Army Corps of Engineers EM 385-1-1 Construction Standard”, Safety 2011 Professional Development Conference Proceedings.
- [9] Wilson, A. M. (2001). Federal procurement policy: Effect on minority-owned business. (Order No. 3003060, GeorgeMasonUniversity). ProQuest Dissertations and Theses, , 166-166 p. Retrieved from [http://ezproxy.indstate.edu:2048/login?url=http://search.proquest.com/docview/250797251?accountid=11592. \(250797251\)](http://ezproxy.indstate.edu:2048/login?url=http://search.proquest.com/docview/250797251?accountid=11592. (250797251)).

## **Biographies**

Scott Arias is an Assistance professor at Eastern Kentucky University (EKU) teaching construction management. His prior positions include Director of Design-Build Construction, Director of Construction, Senior Operations Manager, Force Operations Chief and numerous other project supervision and management positions. His background includes experience in construction management, recruiting, safety and quality control. He is a retired Seabee Builder Senior Chief Petty Officer with duty stations at Naval Mobile Construction Battalion 1 and 7, Navy Recruiting District Michigan, Camp David, National Naval Medical Center Bethesda and Central Command Bahrain. With certifications in safety, construction, scheduling, environmental and quality control he is an expert in both industry and academia. Dr. Arias may be reached at [Scott.Arias@eku.edu](mailto:Scott.Arias@eku.edu).