

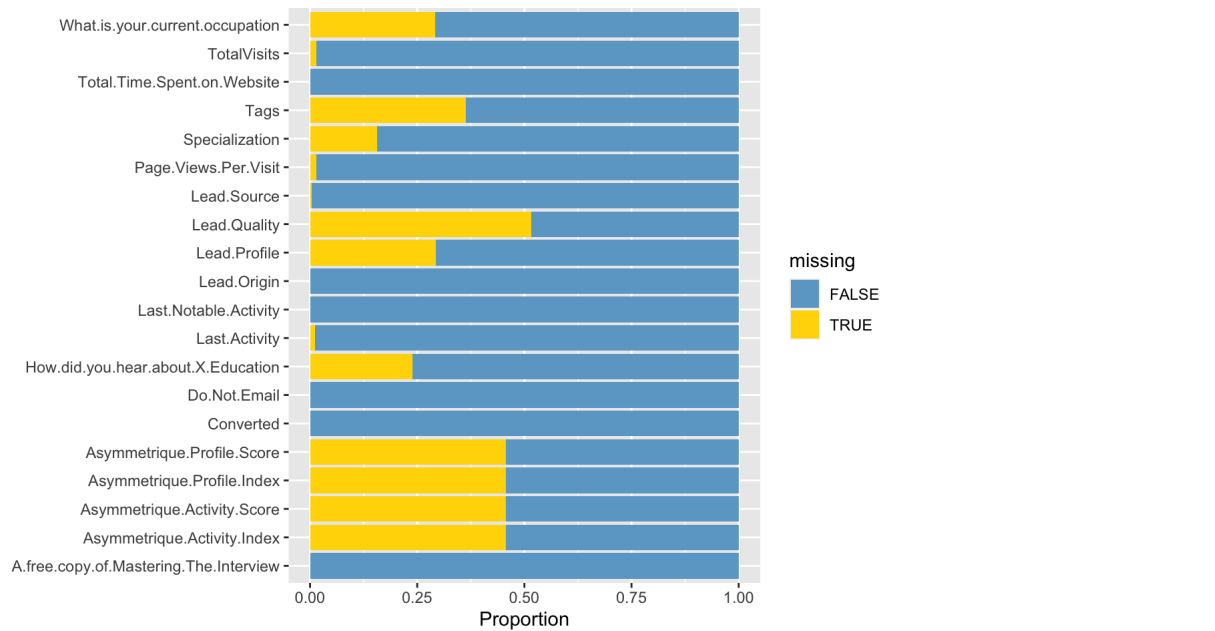
Web Data Analytics

Omkar Sadekar

2023-10-27

```
## 'data.frame': 9240 obs. of 20 variables:
## $ Lead.Origin : chr "API" "API" "Landing Page Submission" "Landing Page Submissio
n" ...
## $ Lead.Source : chr "Olarik Chat" "Organic Search" "Direct Traffic" "Direct Traffi
c" ...
## $ Do.Not.Email : chr "No" "No" "No" "No" ...
## $ Converted : num 0 0 1 0 1 0 1 0 0 0 ...
## $ TotalVisits : num 0 5 2 1 2 0 2 0 2 4 ...
## $ Total.Time.Spent.on.Website : num 0 674 1532 305 1428 ...
## $ Page.Views.Per.Visit : num 0 2.5 2 1 1 0 2 0 2 4 ...
## $ Last.Activity : chr "Page Visited on Website" "Email Opened" "Email Opened" "Unrea
chable" ...
## $ Specialization : chr "Select" "Select" "Business Administration" "Media and Adverti
sing" ...
## $ How.did.you.hear.about.X.Education : chr "Select" "Select" "Select" "Word Of Mouth" ...
## $ What.is.your.current.occupation : chr "Unemployed" "Unemployed" "Student" "Unemployed" ...
## $ Tags : chr "Interested in other courses" "Ringing" "Will revert after rea
ding the email" "Ringing" ...
## $ Lead.Quality : chr "Low in Relevance" NA "Might be" "Not Sure" ...
## $ Lead.Profile : chr "Select" "Select" "Potential Lead" "Select" ...
## $ Asymmetrique.Activity.Index : chr "02.Medium" "02.Medium" "02.Medium" "02.Medium" ...
## $ Asymmetrique.Profile.Index : chr "02.Medium" "02.Medium" "01.High" "01.High" ...
## $ Asymmetrique.Activity.Score : num 15 15 14 13 15 17 14 15 14 13 ...
## $ Asymmetrique.Profile.Score : num 15 15 20 17 18 15 20 15 14 16 ...
## $ A.free.copy.of.Mastering.The.Interview: chr "No" "No" "Yes" "No" ...
## $ Last.Notable.Activity : chr "Modified" "Email Opened" "Email Opened" "Modified" ...
```

```
## [1] 9240 20
```



Lead.Origin	0.0000000
Lead.Source	0.3896104
Do.Not.Email	0.0000000
Converted	0.0000000
TotalVisits	1.4826840
Total.Time.Spent.on.Website	0.0000000
Page.Views.Per.Visit	1.4826840
Last.Activity	1.1147186

x

Specialization	15.5627706
How.did.you.hear.about.X.Education	23.8852814
What.is.your.current.occupation	29.1125541
Tags	36.2878788
Lead.Quality	51.5909091
Lead.Profile	29.3181818
Asymmetrique.Activity.Index	45.6493506
Asymmetrique.Profile.Index	45.6493506
Asymmetrique.Activity.Score	45.6493506
Asymmetrique.Profile.Score	45.6493506
A.free.copy.of.Mastering.The.Interview	0.0000000
Last.Notable.Activity	0.0000000

[1] 6372 14

```
## Rows: 6,372
## Columns: 14
## $ Lead.Origin      <chr> "API", "API", "Landing Page Sub...
## $ Lead.Source      <chr> "Olark Chat", "Organic Search",...
## $ Do.Not.Email     <chr> "No", "No", "No", "No", "No", "...
## $ Converted        <fct> No, No, Yes, No, Yes, Yes, Yes,...
## $ TotalVisits      <dbl> 0, 5, 2, 1, 2, 2, 8, 8, 11, 5, ...
## $ Total.Time.Spent.on.Website <dbl> 0, 674, 1532, 305, 1428, 1640, ...
## $ Page.Views.Per.Visit <dbl> 0.00, 2.50, 2.00, 1.00, 1.00, 2...
## $ Last.Activity    <chr> "Page Visited on Website", "Ema...
## $ Specialization   <chr> "Select", "Select", "Business A...
## $ How.did.you.hear.about.X.Education <chr> "Select", "Select", "Select", "...
## $ What.is.your.current.occupation <chr> "Unemployed", "Unemployed", "St...
## $ Lead.Profile     <chr> "Select", "Select", "Potential ...
## $ A.free.copy.of.Mastering.The.Interview <chr> "No", "No", "Yes", "No", "No", ...
## $ Last.Notable.Activity <chr> "Modified", "Email Opened", "Em...
```

[1] 6098 14

Index	Variable_Name	Variable_Type	Sample_n	Missing_Count	Per_of_Missing	No_of_distinct_values	mean	median
1	Lead.Origin	character	6098	0	0	4	NA	NA
2	Lead.Source	character	6098	0	0	16	NA	NA
3	Do.Not.Email	character	6098	0	0	2	NA	NA
4	Converted	factor	6098	0	0	2	NA	NA
5	TotalVisits	numeric	6098	0	0	11	3.04	3
6	Total.Time.Spent.on.Website	numeric	6098	0	0	1584	524.87	275
7	Page.Views.Per.Visit	numeric	6098	0	0	30	2.36	2
8	Last.Activity	character	6098	0	0	16	NA	NA
9	Specialization	character	6098	0	0	19	NA	NA
10	How.did.you.hear.about.X.Education	character	6098	0	0	10	NA	NA
11	What.is.your.current.occupation	character	6098	0	0	6	NA	NA
12	Lead.Profile	character	6098	0	0	6	NA	NA
13	A.free.copy.of.Mastering.The.Interview	character	6098	0	0	2	NA	NA
14	Last.Notable.Activity	character	6098	0	0	14	NA	NA

Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Lead.Origin	6098						
... API	2071	34%					
... Landing Page Submission	3423	56.1%					
... Lead Add Form	577	9.5%					
... Lead Import	27	0.4%					
Do.Not.Email	6098						
... No	5688	93.3%					
... Yes	410	6.7%					
Converted	6098						
... No	3184	52.2%					
... Yes	2914	47.8%					
TotalVisits	6098	3.035	2.448	0	1	4	10
Total.Time.Spent.on.Website	6098	524.873	562.599	0	27.25	1013	2272
Page.Views.Per.Visit	6098	2.364	1.919	0	1	3	10
What.is.your.current.occupation	6098						
... Businessman	5	0.1%					
... Housewife	9	0.1%					
... Other	13	0.2%					
... Student	189	3.1%					
... Unemployed	5238	85.9%					
... Working Professional	644	10.6%					
Lead.Profile	6098						
... Dual Specialization Student	19	0.3%					
... Lateral Student	19	0.3%					
... Other Leads	458	7.5%					
... Potential Lead	1489	24.4%					
... Select	3879	63.6%					
... Student of SomeSchool	234	3.8%					
A.free.copy.of.Mastering.The.Interview	6098						
... No	4087	67%					
... Yes	2011	33%					

	Vname	Group	TN	nNeg	nZero	nPos	NegInf	PosInf	NA_Value	Per_of_Missing	sum	min	max	mea
3	Page.Views.Per.Visit	All	6098	0	1347	4751	0	0	0	0	14415.54	0	10	2.3
2	Total.Time.Spent.on.Website	All	6098	0	1351	4747	0	0	0	0	3200676.00	0	2272	524.8
1	TotalVisits	All	6098	0	1347	4751	0	0	0	0	18508.00	0	10	3.0

Lead.Origin	Lead.Source	Attribute	Count	sum	mean	median
API	Olark Chat	TotalVisits	881	332	0.3768445	0
API	Organic Search	TotalVisits	310	1296	4.1806452	4

Lead.Origin	Lead.Source	Attribute	Count	sum	mean	median
Landing Page Submission	Direct Traffic	TotalVisits	1726	5964	3.4553882	3
Landing Page Submission	Google	TotalVisits	1218	5357	4.3981938	4
Landing Page Submission	Organic Search	TotalVisits	458	2451	5.3515284	5
API	Referral Sites	TotalVisits	51	232	4.5490196	4
API	Google	TotalVisits	762	2365	3.1036745	3
API	Direct Traffic	TotalVisits	65	294	4.5230769	4
Landing Page Submission	Referral Sites	TotalVisits	13	71	5.4615385	5
Lead Add Form	Reference	TotalVisits	441	86	0.1950113	0
Lead Add Form	Welingak Website	TotalVisits	128	15	0.1171875	0
Lead Add Form	Google	TotalVisits	1	0	0.0000000	0
Lead Import	Facebook	TotalVisits	27	8	0.2962963	0
Lead Add Form	Olark Chat	TotalVisits	2	2	1.0000000	1
Landing Page Submission	Pay per Click Ads	TotalVisits	1	3	3.0000000	3
Landing Page Submission	bing	TotalVisits	2	6	3.0000000	3
API	Social Media	TotalVisits	1	2	2.0000000	2
Landing Page Submission	WeLearn	TotalVisits	1	2	2.0000000	2
Lead Add Form	Live Chat	TotalVisits	2	0	0.0000000	0
API	bing	TotalVisits	1	2	2.0000000	2
Lead Add Form	Click2call	TotalVisits	3	2	0.6666667	0
Landing Page Submission	testone	TotalVisits	1	5	5.0000000	5
Landing Page Submission	Facebook	TotalVisits	1	4	4.0000000	4
Landing Page Submission	Press_Release	TotalVisits	1	6	6.0000000	6
Landing Page Submission	Social Media	TotalVisits	1	3	3.0000000	3

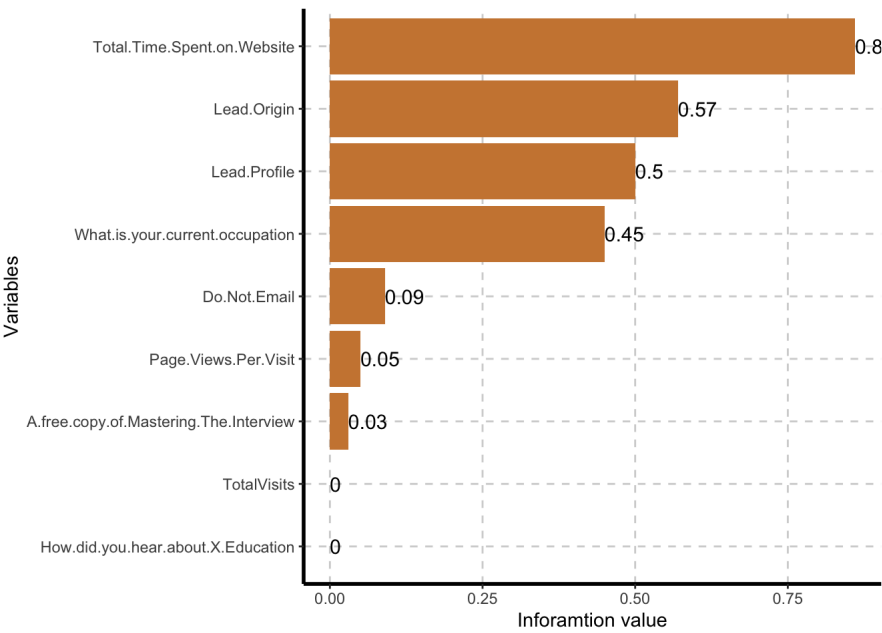
VARIABLE	CATEGORY	Converted/No	Converted/Yes	TOTAL
Lead Origin	API	1159	912	2071
Lead Origin	Landing Page Submission	1970	1453	3423
Lead Origin	Lead Add Form	37	540	577
Lead Origin	Lead Import	18	9	27
Lead Origin	TOTAL	3184	2914	6098
Do_Not_Email	No	2859	2829	5688
Do_Not_Email	Yes	325	85	410
Do_Not_Email	TOTAL	3184	2914	6098
How did you hear about X Education	Advertisements	21	26	47
How did you hear about X Education	Email	9	12	21
How did you hear about X Education	Multiple Sources	63	48	111
How did you hear about X Education	Online Search	272	292	564
How did you hear about X Education	Other	73	66	139
How did you hear about X Education	Select	2484	2207	4691
How did you hear about X Education	SMS	7	5	12
How did you hear about X Education	Social Media	26	23	49
How did you hear about X Education	Student of Somelshool	113	117	230
How did you hear about X Education	Word Of Mouth	116	118	234
How did you hear about X Education	TOTAL	3184	2914	6098
What is your current occupation	Businessman	1	4	5
What is your current occupation	Housewife	0	9	9
What is your current occupation	Other	6	7	13
What is your current occupation	Student	118	71	189
What is your current occupation	Unemployed	3008	2230	5238
What is your current occupation	Working Professional	51	593	644
What is your current occupation	TOTAL	3184	2914	6098
Lead Profile	Dual Specialization Student	0	19	19
Lead Profile	Lateral Student	0	19	19

VARIABLE	CATEGORY	Converted:No	Converted:Yes	TOTAL
Lead:Profile	Other Leads	291	167	458
Lead:Profile	Potential Lead	326	1163	1489
Lead:Profile	Select	2341	1538	3879
Lead:Profile	Student of SomeSchool	226	8	234
Lead:Profile	TOTAL	3184	2914	6098
A.free.copy.of.Mastering.The.Interview	No	1895	2092	4087
A.free.copy.of.Mastering.The.Interview	Yes	1189	822	2011
A.free.copy.of.Mastering.The.Interview	TOTAL	3184	2914	6098

Warning in chisq.test(tb): Chi-squared approximation may be incorrect

Variable	Target	Unique	Chi-squared	p-value	df	IV Value	Cramers V	Degree of Association	Predictive Power
Lead.Origin	Converted	4	538.136	0.000	3	0.57	0.30	Strong	Highly Predictive
Do.Not.Email	Converted	2	127.784	0.000	1	0.09	0.14	Weak	Somewhat Predictive
How.did.you.hear.about.X.Education	Converted	10	9.073	0.431	9	0.00	0.04	Very Weak	Not Predictive
What.is.your.current.occupation	Converted	6	583.465	0.000	5	0.45	0.31	Strong	Highly Predictive
Lead.Profile	Converted	6	901.205	0.000	5	0.50	0.38	Strong	Highly Predictive
A.free.copy.of.Mastering.The.Interview	Converted	2	57.024	0.000	1	0.03	0.10	Weak	Not Predictive
TotalVisits	Converted	6	11.992	0.035	5	0.00	0.04	Very Weak	Not Predictive
Total.Time.Spent.on.Website	Converted	8	1166.617	0.000	7	0.86	0.44	Strong	Highly Predictive
Page.Views.Per.Visit	Converted	7	78.101	0.000	6	0.05	0.11	Weak	Not Predictive

Warning in chisq.test(tb): Chi-squared approximation may be incorrect

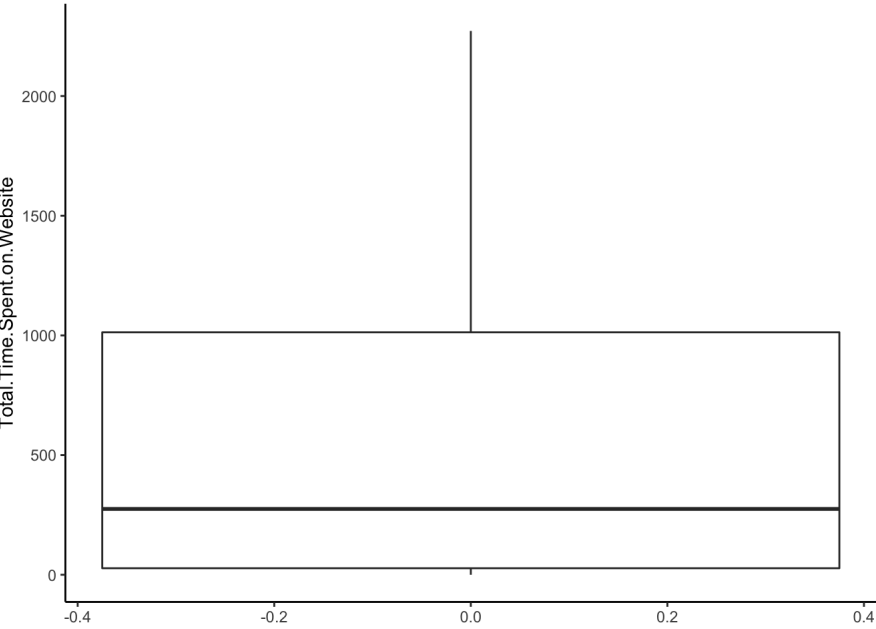


Warning in chisq.test(tb): Chi-squared approximation may be incorrect

Variable	Target	Unique	Chi-squared	p-value	df	IV Value	Cramers V	Degree of Association	Predictive Power
Lead.Origin	Converted	4	538.136	0.000	3	0.57	0.30	Strong	Highly Predictive
Do.Not.Email	Converted	2	127.784	0.000	1	0.09	0.14	Weak	Somewhat Predictive
How.did.you.hear.about.X.Education	Converted	10	9.073	0.431	9	0.00	0.04	Very Weak	Not Predictive
What.is.your.current.occupation	Converted	6	583.465	0.000	5	0.45	0.31	Strong	Highly Predictive
Lead.Profile	Converted	6	901.205	0.000	5	0.50	0.38	Strong	Highly Predictive
A.free.copy.of.Mastering.The.Interview	Converted	2	57.024	0.000	1	0.03	0.10	Weak	Not Predictive
TotalVisits	Converted	6	11.992	0.035	5	0.00	0.04	Very Weak	Not Predictive

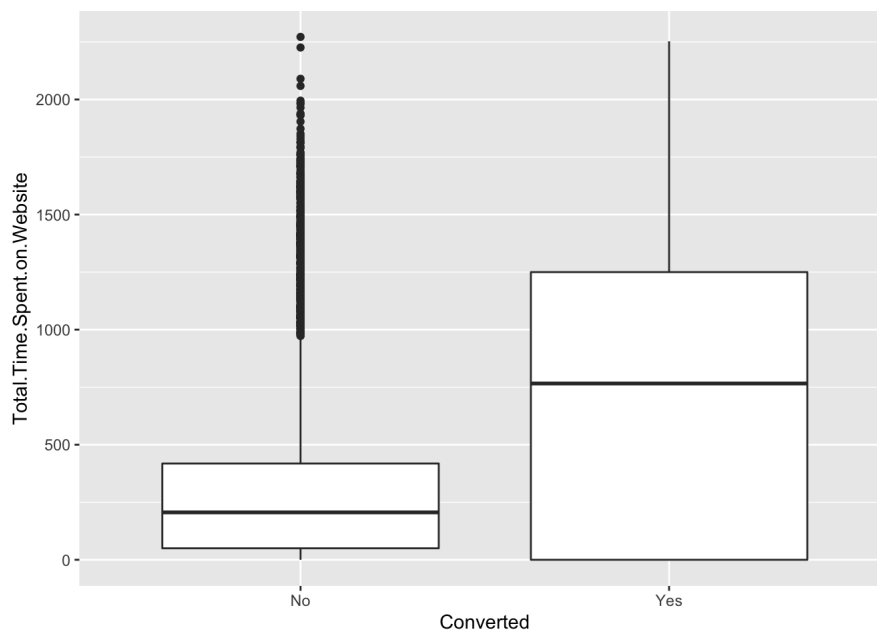
Variable	Target	Unique	Chi-squared	p-value	df	IV Value	Cramers V	Degree of Association	Predictive Power
Total.Time.Spent.on.Website	Converted	8	1166.617	0.000	7	0.86	0.44	Strong	Highly Predictive
Page.Views.Per.Visit	Converted	7	78.101	0.000	6	0.05	0.11	Weak	Not Predictive

```
## [1] "Lead.Origin"
## [2] "Lead.Source"
## [3] "Do.Not.Email"
## [4] "Converted"
## [5] "TotalVisits"
## [6] "Total.Time.Spent.on.Website"
## [7] "Page.Views.Per.Visit"
## [8] "Last.Activity"
## [9] "Specialization"
## [10] "How.did.you.hear.about.X.Education"
## [11] "What.is.your.current.occupation"
## [12] "Lead.Profile"
## [13] "A.free.copy.of.Mastering.The.Interview"
## [14] "Last.Notable.Activity"
```



```
##
## One Sample t-test
##
## data: Leads_df$Total.Time.Spent.on.Website
## t = 38.153, df = 6097, p-value = 1
## alternative hypothesis: true mean is less than 250
## 95 percent confidence interval:
##      -Inf 536.7253
## sample estimates:
## mean of x
## 524.8731
```

```
##
## One Sample t-test
##
## data: Leads_df$Total.Time.Spent.on.Website
## t = 38.153, df = 6097, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 250
## 95 percent confidence interval:
## 510.7497 538.9965
## sample estimates:
## mean of x
## 524.8731
```



```
##
## Two Sample t-test
##
## data: Converted$Total.Time.Spent.on.Website and NotConverted$Total.Time.Spent.on.Website
## t = 16.562, df = 1998, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 437.3941
## sample estimates:
## mean of x mean of y
## 740.057 342.195
```

```
##
## Two Sample t-test
##
## data: Converted$Total.Time.Spent.on.Website and NotConverted$Total.Time.Spent.on.Website
## t = 16.562, df = 1998, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 350.7499 444.9741
## sample estimates:
## mean of x mean of y
## 740.057 342.195
```

```
##
## Pearson's Chi-squared test
##
## data: Leads_df$Lead.Origin and Leads_df$Converted
## X-squared = 538.14, df = 3, p-value < 2.2e-16
```

```
##
## Leads_df$Lead.Origin      Leads_df$Converted
##      API      1081.34864      989.65136
## Landing Page Submission 1787.27976 1635.72024
## Lead Add Form      301.27386 275.72614
## Lead Import      14.09774 12.90226
```

```
##
##      API Landing Page Submission      Lead Add Form
##      2071      3423      577
##      Lead Import
##      27
```

```
##
## Chi-squared test for given probabilities
##
## data: deg_count
## X-squared = 267.74, df = 3, p-value < 2.2e-16
```

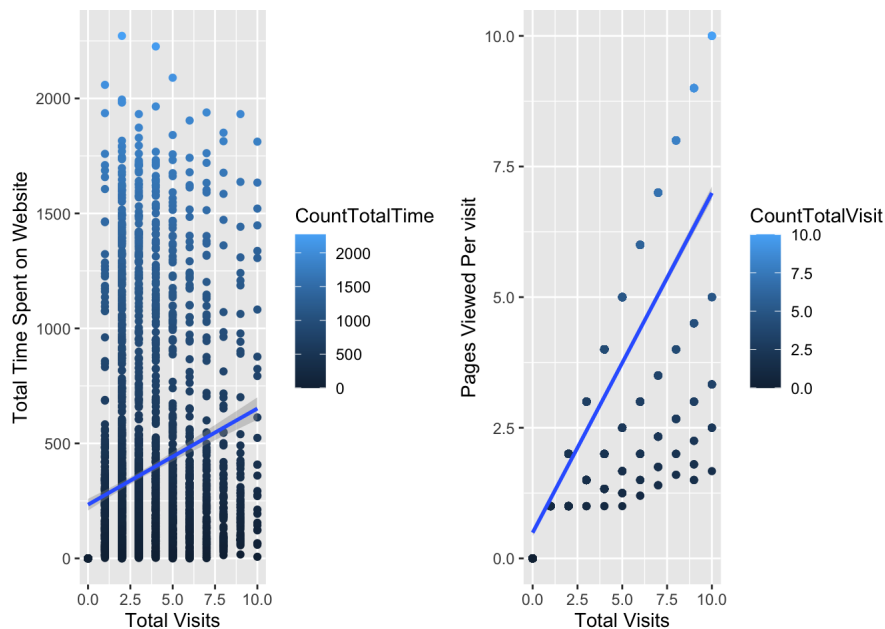
##	API Landing Page Submission	Lead Add Form
##	2012.34	548.82
##	Lead Import	
##	304.90	

##	Lead.Origin	Lead.Source	Do.Not.Email	Converted	TotalVisits
## 1:	API	Olark Chat	No	No	0
## 2:	API	Organic Search	No	No	5
## 3: Landing Page Submission	Direct Traffic		No	Yes	2
## 4: Landing Page Submission	Direct Traffic		No	No	1
## 5: Landing Page Submission	Google		No	Yes	2
## ---					
## 6094: Landing Page Submission	Direct Traffic		No	Yes	5
## 6095: Landing Page Submission	Direct Traffic		Yes	Yes	8
## 6096: Landing Page Submission	Direct Traffic		No	No	2
## 6097: Landing Page Submission	Direct Traffic		Yes	No	2
## 6098: Landing Page Submission	Direct Traffic		No	Yes	6
##	Total.Time.Spent.on.Website	Page.Views.Per.Visit			Last.Activity
## 1:	0	0.00			Page Visited on Website
## 2:	674	2.50			Email Opened
## 3:	1532	2.00			Email Opened
## 4:	305	1.00			Unreachable
## 5:	1428	1.00			Converted to Lead
## ---					
## 6094:	210	2.50			SMS Sent
## 6095:	1845	2.67			Email Marked Spam
## 6096:	238	2.00			SMS Sent
## 6097:	199	2.00			SMS Sent
## 6098:	1279	3.00			SMS Sent
##	Specialization	How.did.you.hear.about.X.Education			
## 1:	Select			Select	
## 2:	Select			Select	
## 3: Business Administration				Select	
## 4: Media and Advertising		Word Of Mouth			
## 5:	Select			Other	
## ---					
## 6094: Business Administration				Select	
## 6095: IT Projects Management				Select	
## 6096: Media and Advertising				Select	
## 6097: Business Administration				Select	
## 6098: Supply Chain Management				Select	
##	What.is.your.current.occupation	Lead.Profile			
## 1:	Unemployed	Select			
## 2:	Unemployed	Select			
## 3:	Student	Potential Lead			
## 4:	Unemployed	Select			
## 5:	Unemployed	Select			
## ---					
## 6094:	Unemployed	Potential Lead			
## 6095:	Unemployed	Potential Lead			
## 6096:	Unemployed	Potential Lead			
## 6097:	Unemployed	Potential Lead			
## 6098:	Unemployed	Potential Lead			
##	A.free.copy.of.Mastering.The.Interview	Last.Notable.Activity	Converted_No		
## 1:	No	Modified	1		
## 2:	No	Email Opened	1		
## 3:	Yes	Email Opened	0		
## 4:	No	Modified	1		
## 5:	No	Modified	0		
## ---					
## 6094:	No	Modified	0		
## 6095:	No	Email Marked Spam	0		
## 6096:	Yes	SMS Sent	1		
## 6097:	Yes	SMS Sent	1		
## 6098:	Yes	Modified	0		
##	Converted_Yes				
## 1:	0				
## 2:	0				
## 3:	1				
## 4:	0				
## 5:	1				
## ---					
## 6094:	1				
## 6095:	1				
## 6096:	0				
## 6097:	0				
## 6098:	1				

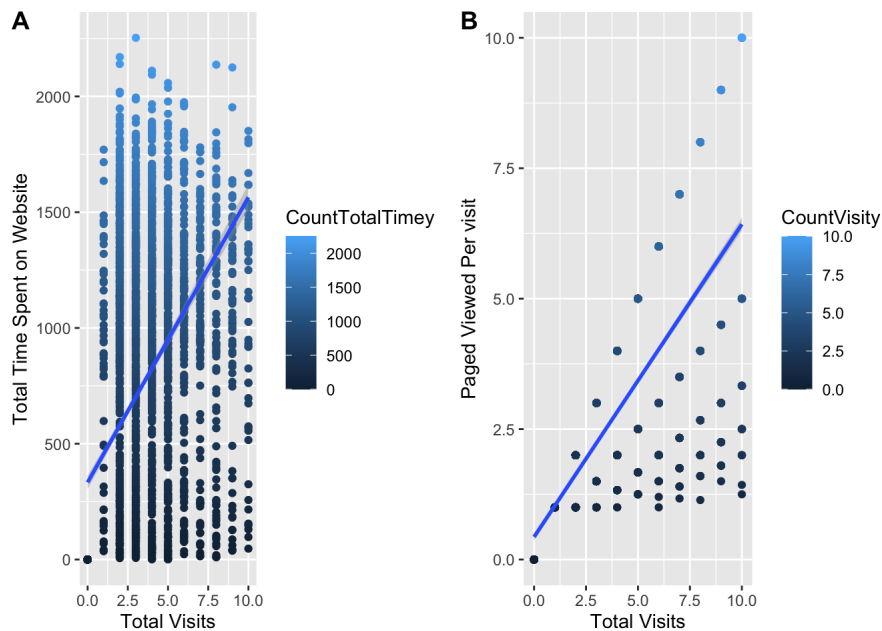

```
## TotalVisits Total.Time.Spent.on.Website Page.Views.Per.Visit
## Min. : 0.000 Min. : 0.0 Min. : 0.000
## 1st Qu.: 2.000 1st Qu.: 50.0 1st Qu.: 1.250
## Median : 3.000 Median : 206.0 Median : 2.000
## Mean : 3.062 Mean : 361.6 Mean : 2.483
## 3rd Qu.: 4.000 3rd Qu.: 418.2 3rd Qu.: 3.000
## Max. :10.000 Max. :2272.0 Max. :10.000
```

```
## TotalVisits Total.Time.Spent.on.Website Page.Views.Per.Visit
## Min. : 0.000 Min. : 0.0 Min. : 0.000
## 1st Qu.: 0.000 1st Qu.: 0.0 1st Qu.: 0.000
## Median : 3.000 Median : 766.0 Median : 2.000
## Mean : 3.006 Mean : 703.3 Mean : 2.234
## 3rd Qu.: 5.000 3rd Qu.:1250.0 3rd Qu.: 3.000
## Max. :10.000 Max. :2253.0 Max. :10.000
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

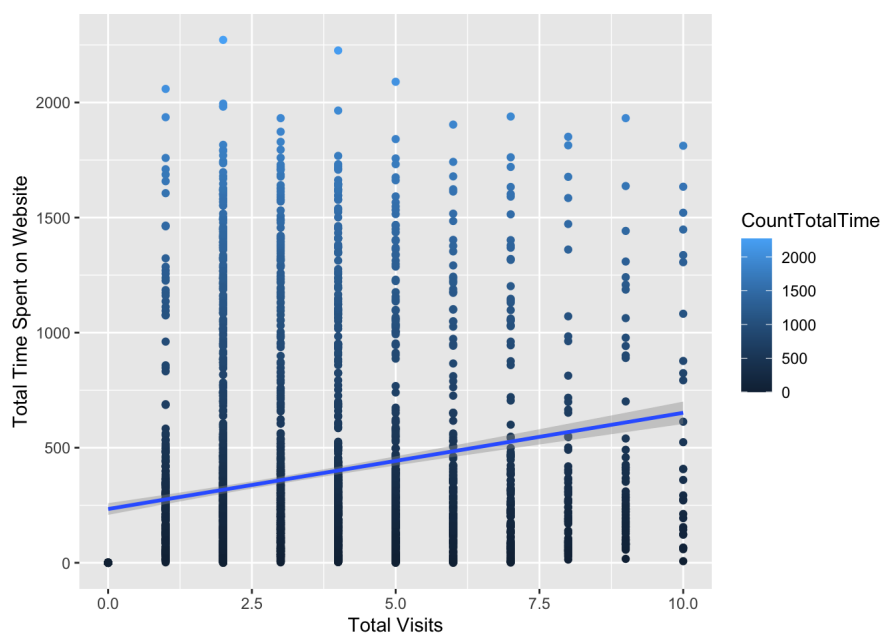


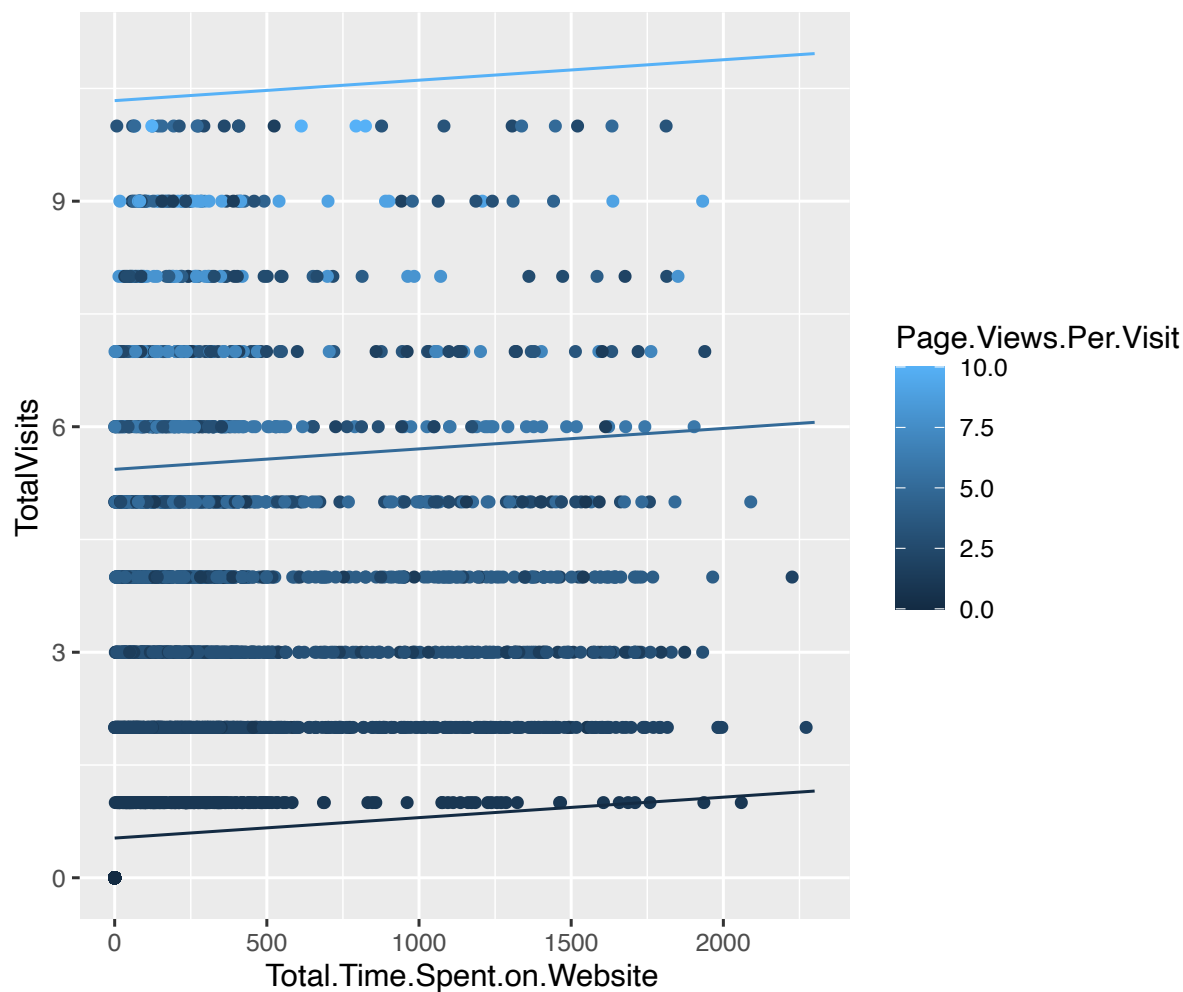
```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```



```
##
## Call:
## lm(formula = TotalVisits ~ Total.Time.Spent.on.Website + Page.Views.Per.Visit,
##     data = S_no)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1079 -0.5669 -0.5279 -0.4605  7.6913
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.279e-01  4.230e-02  12.48 < 2e-16 ***
## Total.Time.Spent.on.Website 2.720e-04  5.485e-05   4.96 7.43e-07 ***
## Page.Views.Per.Visit    9.810e-01  1.326e-02  73.97 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.355 on 3181 degrees of freedom
## Multiple R-squared:  0.6492, Adjusted R-squared:  0.6489
## F-statistic: 2943 on 2 and 3181 DF, p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula 'y ~ x'
```





```
##
## Call:
## lm(formula = TotalVisits ~ Total.Time.Spent.on.Website + Page.Views.Per.Visit,
##     data = S_yes)
##
## Residuals:
##      Min       10   Median       30      Max
## -1.6996 -0.8847 -0.4719 -0.2560  8.2222
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   4.719e-01  4.768e-02   9.897  <2e-16 ***
## Total.Time.Spent.on.Website 6.305e-04  5.512e-05  11.438  <2e-16 ***
## Page.Views.Per.Visit    9.357e-01  1.719e-02  54.442  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.572 on 2911 degrees of freedom
## Multiple R-squared:  0.6384, Adjusted R-squared:  0.6382
## F-statistic: 2570 on 2 and 2911 DF, p-value: < 2.2e-16
```

```
## `geom_smooth()` using formula 'y ~ x'
```

