

Exploratory Data Analysis

Capstone Step 5

Choices Regarding Data Cleaning/Transformation

During prototyping I observed a few high level improvements possible via treating the raw data. Specific choices are annotated in screenshots on the subsequent slides.

1. Location file is entirely redundant and will be descoped.
2. A number of columns are redundant and will be removed after verifying the other locations offer equivalent data
3. A number of columns can be joined/consolidated
4. A number of columns can be converted from varchar to numerical type

Choices Regarding Pipeline Automation

The prototyping phase helped form the following approach to automation in order to achieve the optimal reliability:

1. To handle embedded commas within csv columns, use Linux csvkit's csvcut tool from Python (cut and awk both have tremendous difficulty).
- 2 . Use sqlalchemy create_engine, pandas read_csv and dataframe.to_sql instead of looping through each row to execute mysql_connector_python insert statement. The latter fails in cases of NULL values.
3. Do not declare any columns NOT NULL, and err on the side of caution with strings (TEXT instead of VARCHAR).
4. Delete files after processing to keep to manageable level.

Pipeline Script: transformdata.py

```
transformdata.py
1  """
2  Script to transform raw csv files into clean final format and load to mysql
3  """
4
5  import mysql.connector
6  from mysql.connector import errorcode
7  from pathlib import Path
8  import os
9  import glob
10 import pandas as pd
11 from sqlalchemy import create_engine, types
12 import csv
13
14 def get_db_connection():
15     """
16     2.1 Setup database connection
17     In order to make a query against the database table, we need to first connect to it. A connection
18     can be established only when the user provides the proper target host, port, and user
19     credentials
20     """
21     connection = None
22     try:
23         connection = mysql.connector.connect(user='root', password='', host='127.0.0.1', port='3306')
24     except Exception as error:
25         print("Error while connecting to database for job tracker", error)
26     return connection
27
28 def initialize_database():
29     cnx = get_db_connection()
30     cursor = cnx.cursor()
31     try:
32         cursor.execute("DROP DATABASE IF EXISTS {}".format(DB_NAME))
33         cursor.execute("CREATE DATABASE {}".format(DB_NAME))
34     except mysql.connector.Error as err:
35         print("Failed creating database: {}".format(err))
36         exit(1)
37     try:
38         cursor.execute("USE {}".format(DB_NAME))
39     except mysql.connector.Error as err:
40         print("Database {} does not exists.".format(DB_NAME))
41         if err.errno == errorcode.ER_BAD_DB_ERROR:
42             create_database(cursor)
43         print("Database {} created successfully.".format(DB_NAME))
44         cnx.database = DB_NAME
45     else:
46         print(err)
47         exit(1)
48
49 TABLES = {}
50 TABLES['details'] = (
51     "CREATE TABLE details ("
52     "    BEGIN YEARMONTH VARCHAR(6),"
53     "    BEGIN DAY VARCHAR(2),"
54     "    BEGIN TIME VARCHAR(4),"
55     "    END YEARMONTH VARCHAR(6),"
56     "    END DAY VARCHAR(2),"
57     "    END TIME VARCHAR(4),"
58     "    EPISODE ID INT,"
59     "    EVENT ID INT,"
60     "    STATE TEXT,"
```

```
transformdata.py
59
60 TABLES['fatalities'] = (
61     "CREATE TABLE fatalities ("
62     "    FATALITY ID INT,"
63     "    EVENT ID INT,"
64     "    FATALITY TYPE VARCHAR(3),"
65     "    FATALITY DATE VARCHAR(10),"
66     "    FATALITY AGE INT DEFAULT NULL,"
67     "    FATALITY SEX ENUM('M','F'),"
68     "    FATALITY LOCATION TEXT,"
69     "    EVENT YEARMONTH VARCHAR(6)"
70     ")")
71
72 for table_name in TABLES:
73     table_description = TABLES[table_name]
74     try:
75         print("Creating table {}: {}".format(table_name, and=""))
76         cursor.execute(table_description)
77     except mysql.connector.Error as err:
78         if err.errno == errorcode.ER_TABLE_EXISTS_ERROR:
79             print("already exists.")
80         else:
81             print(err.msg)
82     else:
83         print("OK")
84
85 return cnx
86
87 def load_csv(filepath, table):
88     df = pd.read_csv(filepath, sep=',')
89     df.to_sql(table, cnx, engine, index=False, if_exists='append')
90
91 def cut_details(filepath):
92     #details needs cols 11 12 18 20 51 removed
93     os.system('csvcut -columns=1,2,3,4,5,6,7,8,9,10,13,14,15,16,17,19,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50 * ' + filepath + " > details" + filepath[-10:])
94
95 def cut_fatal(filepath):
96     #fatalities needs cols 1 2 3 removed
97     os.system('csvcut -columns=4,5,6,7,8,9,10,11 * ' + filepath + " > fatalities" + filepath[-10:])
98
99 if __name__ == "__main__":
100     print("**** Start of data transform script ****\n")
101     DB_NAME = "Capstone"
102     cnx = initialize_database()
103     engine = create_engine('mysql+psymysql://root@localhost/Capstone')
104     detlist = sorted(glob.glob("/home/conner/Capstone/data/unzipped/storm_details.*"))
105     fatlist = sorted(glob.glob("/home/conner/Capstone/data/unzipped/storm_fatalities.*"))
106     for d in detlist:
107         print("Cutting " + d)
108         cut_details(d)
109         print("Loading " + d[-25:])
110         load_csv(d[-25:], 'details')
111         os.system('rm ' + d[-25:])
112     for f in fatlist:
113         print("Cutting " + f)
114         cut_fatal(f)
115         print("Loading " + f[-20:])
116         load_csv(f[-20:], 'fatalities')
117         os.system('rm ' + f[-20:])
118     print("**** End of script ****")
```

Details

	A	R	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V			
BEGIN_YEAR	YEAR	BEGIN_YEAR	YEAR	END_YEAR	END_YEAR	END_YEAR	PROJID	ID	EVENT_ID	STATE	DATE	EVENT	MONTH	NAME	EVENT_TYPE	CZ_TYPE	CZ_FIPS	CZ_NAME	WQ	BEGIN_DATE	TIME	END_DATE	TIME	INJURIES_DIRECT	INJURIES_INDIRECT
23	207	202004	23	228	202004	23	47469	887009	MISSISSIPPI	28	2020	April	Tornado	C	5	AMITE	LIX	23-APR-20 02:07:00	ST-6	3-APR-20 02:00:00		0	0		
3	950	202008	4	951	202008	4	48216	904084	PENNSYLVANIA	42	2020	August	Tornado	C	17	BUCKS	PHI	04-AUG-20 09:55:00	ST-5	4-AUG-20 09:00:00		0	0		
6	529	202002	6	529	202002	6	45803	875896	GEORGIA	13	2020	February	Tornado	C	15	BARTOW	FFC	06-FEB-20 06:20:00	ST-5	6-FEB-20 06:00:00		0	0		
6	1055	202002	6	1100	202002	6	45216	871938	NORTH CAROLINA	37	2020	February	Tornado	C	45	CLEVELAND	GRF	06-FEB-20 10:10:00	ST-5	6-FEB-20 10:00:00		0	0		
13	1715	202007	25	1315	202007	25	50353	913730	TEXAS	48	2020	July	Hurricane	Z	256	COASTAL WILLACY	BSC	25-JUL-20 13:49:00	ST-6	5-JUL-20 13:00:00		0	0		
25	1400	202007	25	1800	202007	25	51466	912708	TEXAS	48	2020	July	Hurricane	Z	442	KLEBERG COUNTY ISLANDS	CRP	25-JUL-20 14:00:00	ST-6	5-JUL-20 14:00:00		0	0		
26	1800	202008	27	600	202008	27	52434	918106	TEXAS	48	2020	August	Hurricane	Z	216	ORANGE	LCH	26-AUG-20 13:00:00	ST-6	7-AUG-20 00:00:00		0	0		
25	1700	202005	25	2000	202005	25	47310	885808	WEST VIRGINIA	54	2020	May	Flash Flood	C	101	WEBSTER	RLX	25-MAY-20 00:00:00	ST-6	5-MAY-20 00:00:00		0	0		
13	30	202004	13	330	202004	13	45523	880416	VIRGINIA	51	2020	April	Flood	C	51	DICKENSON	RLX	13-APR-20 08:30:00	ST-5	3-APR-20 08:30:00		0	0		
28	1830	202003	28	1930	202003	28	45618	874566	OHIO	39	2020	March	Flash Flood	C	37	DAIRY	ILN	28-MAR-20 05:30:00	ST-5	3-MAR-20 05:30:00		0	0		
28	1530	202003	28	1530	202003	28	45617	874465	OHIO	39	2020	March	Flood	C	89	LICKING	ILN	28-MAR-20 05:30:00	ST-5	3-MAR-20 05:30:00		0	0		
13	1900	202004	14	545	202004	14	46730	883021	OHIO	39	2020	April	Flood	C	167	WASHINGTON	RLX	13-APR-20 05:00:00	ST-5	4-APR-20 05:45:00		0	0		
9	1401	202009	9	1615	202009	9	52853	920637	PUERTO RICO	99	2020	September	Flash Flood	C	127	SAN JUAN	SJU	09-SEP-20 01:00:00	ST-4	9-SEP-20 01:15:00		0	0		
9	1405	202009	9	1615	202009	9	52853	920638	PUERTO RICO	99	2020	September	Flash Flood	C	127	SAN JUAN	SJU	09-SEP-20 14:05:00	ST-4	9-SEP-20 16:15:00		0	0		
6	1605	202003	6	1606	202003	6	46986	863678	SOUTH CAROLINA	45	2020	March	Strong Wind	Z	47	JASPER	CHE	06-MAR-20 16:00:00	ST-5	6-MAR-20 16:00:00		0	0		
31	1451	202003	31	1452	202003	31	46987	863679	SOUTH CAROLINA	45	2020	March	Thunderstorm Wind	C	5	ALLENDALE	CHE	31-MAR-20 14:00:00	ST-5	31-MAR-20 14:00:00		0	0		
6	1230	202002	6	1330	202002	6	44988	870686	FLORIDA	12	2020	February	High Wind	Z	136	EASTERN ALACHUA	JAX	06-FEB-20 12:00:00	ST-5	6-FEB-20 13:00:00		0	0		
13	751	202004	13	816	202004	13	47487	887356	ATLANTIC SOUTH	87	2020	April	Marine Thunderstorm Win	Z	350	S Santee R to Edisto Beach SC Out 20NM	CHE	13-APR-20 07:50:00	ST-5	3-APR-20 08:14:00		0	0		
13	804	202004	13	809	202004	13	47487	887357	ATLANTIC SOUTH	87	2020	April	Marine Thunderstorm Win	Z	350	S Santee R to Edisto Beach SC Out 20NM	CHE	13-APR-20 08:04:00	ST-5	3-APR-20 08:08:00		0	0		
15	341	202004	15	341	202004	15	47255	885395	ATLANTIC NORTH	88	2020	April	Marine Strong Wind	Z	634	CHESAPEAKE BAY FROM LITTLE CREEK, VA, TO CAPE HENRY, VA, INCLUDING THE CHESAPEAKE BAY BRIDGE TUNNEL	LAK	15-APR-20 03:41:00	ST-5	3-APR-20 03:41:00		0	0		
13	800	202004	13	1225	202004	13	46438	879963	LAKE SUPERIOR	92	2020	April	Marine High Wind	Z	266	MANTOU ISLAND TO MARQUETTE MI AND W OF GRAND MARAIS MI TO US/CANADIAN BORDER BEYOND 5NM OF SHORE	MOT	13-APR-20 08:00:00	ST-5	3-APR-20 12:25:00		0	0		

dt dt i i v i v v i v v i i i

Convert 6 columns to 2 columns: begin/end datetime

Redundant (delete after verifying equivalence)

Redundant (delete after verifying equivalence)

DATATYPES
v - varchar
dt - datetime
i - int

U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ							
INJURIES_DIRECT	INJURIES_INDIRECT	DEATHS_DIRECT	DEATHS_INDIRECT	DAMAGE_PROPERTY	DAMAGE_CROPS	SOURCE	MAGNITUDE	MAGNITUDE	TYPE	FLOOD	CAUSE	CATEGORY	TOR_F	SCALE	TOR_LENGTH	TOR_WIDTH	TOR_OTHER	WFO_TOR	OTHER_CZ	STATE	TOR_OTHER	CZ_RPS	TOR_OTHER_CZ	CZ_NAME	BEGIN_RANGE	BEGIN_ANGLE	BEGIN_LOCATION	END_RANGE	END_ANGLE
2	0	0	0	100.00K	0.00K	Emergency Manager						EF2	18.3	1760	LIX	MS	113		W				PIKE	4	W	EAST FORK	11	E	
3	0	0	0			NWS Storm Survey						EF2	0.06	500	PH	PA	101		W				PHILADELPHIA	2	W	CORNWELLS HGTS	2	W	
4	0	0	0	10.00K		NWS Storm Survey						EF0	0.33	200	FFC	GA	129		WNW				GORDON	2	WNW	FOLSON	2	NW	
5	0	0	0	750.00K	0.00K	NWS Storm Survey						EF2	5.14	150	GSP	NC	71		ESE				GASTON	1	ESE	ARCHDALE	4	SSE	
6	0	0	0	127.60K		Booy				1																			
7	0	0	0	50.00K	0.00K	Mesonet				1																			
8	0	0	0	0.10B	0.00K	Emergency Manager				2																			
9	0	0	0	3.00K	0.00K	Department of Highways			Heavy Rain															1	NNW	ERBACON	2	E	
10	0	0	0	5.00K	0.00K	911 Call Center			Heavy Rain															1	E	OSBORNS GAP	2	NE	
11	0	0	0	0.00K	0.00K	State Official			Heavy Rain															1	E	GREENVILLE	2	S	
12	0	0	0	0.00K	0.00K	Law Enforcement			Heavy Rain															2	ESE	PATASKALA ZELINE ARPT	2	NW	
13	0	0	0	2.00K	0.00K	River/Stream Gage			Heavy Rain															1	N	MACKSBURG	1	NE	
14	0	0	0	0.00K	0.00K	911 Call Center			Heavy Rain															4	NW	SAN JUAN	3	NW	
15	0	0	0	0.00K	0.00K	911 Call Center			Heavy Rain															4	NW	SAN JUAN	4	NW	
16	0	0	0	10.00K	0.00K	Law Enforcement	35	EG																					
17	0	0	0	0.00K	0.00K	911 Call Center	50	EG																2	WNW	ALLEDALE OSWALD ARP	2	WNW	
18	0	0	0	5.00K	0.00K	Mesonet	41	MG																1	WSW	FOLLY BEACH	1	WSW	
19	0	0	0	0.00K	0.00K	Mesonet	49	MG																1	E	SULLIVANS ISLAND	1	E	
20	0	0	0	0.00K	0.00K	Mesonet	42	MG																0	N	CHESAPEAKE BAY BRIDGE TUNNEL 3RD ISLAND	0	N	
21	0	0	0	1.00K	0.00K	Mesonet	36	MG																0	N	STANNARD ROCK LIGHT	0	N	
22	0	0	0	0.00K	0.00K	C-MAN Station	51	MG																0	N		0	N	

Convert to integer by multiplying out based on suffix (e.g. K = 1000x)

i i i i i i v i v v i v f i v i v v i v

DATATYPES:
i - int
v - varchar
f - float

Details (cont)

	AQ	AR	AS	AT	AU	AV	AW	AX	AY
	END_AZIMUTH	END_LOCATION	BEGIN_LAT	BEGIN_LON	END_LAT	END_LON	EPISODE	N_EVENT	NAT_DATA_SOURCE
2	E	MARS HILL	31.2047	-90.7432	31.2705	-90.446	A strong co/A very large	25V	
3	W	CORNWELLS HGTS	40.0815	-74.9592	40.0822	-74.9599	Tropical Storm tornado	25V	
4	NW	FOL SOM	34.3951	-84.8631	34.3968	-84.8576	A line of thru a National	25V	
5	SSE	KINGS MTN	35.177	-81.413	35.196	-81.325	Unusually h/NWS storm	25V	
6							Hurricane HWHurricane H	25V	
7							Hurricane HWTCOON site	25V	
8							A tropical wt there was	25V	
9	E	ERBACON	38.5367	-80.5887	38.5186	-80.5378	Scattered apCounty Road	25V	
10	NE	BARTLICK	37.1988	-82.5269	37.2685	-82.3034	An intense WAt the inter	25V	
11	S	WEIMERS MILL	40.1	-84.62	40.0942	-84.6748	ThunderstormHigh water	25V	
12	NW	KIRKERSVILLE	39.97	-82.62	39.969	-82.6201	Scattered thHigh water	25V	
13	NE	MACKSBURG	39.6391	-81.4691	39.6392	-81.4539	An intense WOn the West	25V	
14	NW	SAN JUAN	18.4148	-66.0901	18.4152	-66.0878	An upper-levFlooding rep	25V	
15	NW	SAN JUAN	18.414	-66.1016	18.4139	-66.101	An upper-levFlooding rep	25V	
16							Strong gradLow enforce	25V	
17	WNW	ALLENDALE OSWALD ARP	33.01	-81.4	33.01	-81.4	An area of h/Alendale C	25V	
18							Very strong Strong grad	25V	
19	WSW	FOLLY BEACH	32.65	-79.94	32.65	-79.94	A severe quThe Weather	25V	
20	E	SULLIVANS ISLAND	32.76	-79.82	32.76	-79.82	A severe quThe Weather	25V	
21	N	CHESAPEAKE BAY BRIDGE TUNNEL 3RD ISLAND	37.04	-76.08	37.04	-76.08	Scattered shWind gust	25V	
22	N	STANNARD ROCK LIGHT	47.18	-87.23	47.18	-87.23	A potent stoThe Stannard	25V	
23	V	V	F	F	F	F	V	V	Delete

Fatalities

	A	B	C	D	E	F	G	H	I	J	K
1	FAT_YEARMONTH	FAT_DAY	FAT_TIME	FATALITY_ID	EVENT_ID	FATALITY_TYPE	FATALITY_DATE	FATALITY_AGE	FATALITY_SEX	FATALITY_LOCATION	EVENT_YEARMONTH
2	202106	9	0	42960	953511	D	06/09/2021 00:00:00	70	M	Golfing	202106
3	202107	20	0	43206	961309	I	07/20/2021 00:00:00	32	M	Other	202107
4	202107	20	0	43207	961309	I	07/20/2021 00:00:00	31	M	Other	202107
5	202107	20	0	43208	961309	I	07/20/2021 00:00:00		M	Other	202107
6	202107	27	0	44279	964033	I	07/27/2021 00:00:00		F	Other	202107
7	202103	28	0	42898	951050	D	03/28/2021 00:00:00	61	F	In Water	202103
8	202104	24	0	42962	954336	D	04/24/2021 00:00:00	2		Permanent Home	202104
9	202106	10	0	42963	954408	D	06/10/2021 00:00:00		M	In Water	202106
10	202106	14	0	44478	970299	D	06/14/2021 00:00:00	23	M	Outside/Open Areas	202106
11	202106	13	0	44482	970319	D	06/13/2021 00:00:00	32	F	Outside/Open Areas	202106
12	202106	13	0	44483	970319	D	06/13/2021 00:00:00	42	M	Outside/Open Areas	202106
13	202106	14	0	44484	970319	D	06/13/2021 00:00:00	29	M	Outside/Open Areas	202106
14	202106	15	0	44485	970319	D	06/15/2021 00:00:00	34	F	Outside/Open Areas	202106
15	202106	17	0	44487	970319	D	06/17/2021 00:00:00	28	M	Outside/Open Areas	202106
16	202106	20	0	44488	970319	D	06/20/2021 00:00:00	19	M	Outside/Open Areas	202106
17	202106	20	0	44489	970319	D	06/20/2021 00:00:00	42	M	Outside/Open Areas	202106
18	202106	15	0	44490	970319	D	06/15/2021 00:00:00	35	M	Outside/Open Areas	202106
19	202106	17	0	44491	970319	D	06/17/2021 00:00:00	36	M	Outside/Open Areas	202106
20	202106	19	0	44492	970319	D	06/19/2021 00:00:00	54	M	Outside/Open Areas	202106
21	202106	13	0	44479	970320	D	06/13/2021 00:00:00	35	F	Outside/Open Areas	202106
22	202106	12	0	44480	970320	D	06/12/2021 00:00:00	22	M	Outside/Open Areas	202106
23	202106	15	0	44481	970320	D	06/15/2021 00:00:00	37	M	Outside/Open Areas	202106
24	202106	2	0	43761	970487	D	06/02/2021 00:00:00		M	Outside/Open Areas	202106
25	202106	2	0	43762	970487	D	06/02/2021 00:00:00	20	M	Outside/Open Areas	202106
26	202106	7	0	43770	970488	D	06/07/2021 00:00:00	40	M	Outside/Open Areas	202106
27	202106	6	0	43771	970490	D	06/06/2021 00:00:00	47	M	Outside/Open Areas	202106
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											
38											
39											
40											
41											
42											
43											
44											
45											
46											
47											
48											
49											
50											
51											
52											
53											
54											

Redundant
(Will remove after
verifying equivalent
to FATALITY_DATE)

I I V DT I V V I

Datatypes:
I - Int
V - Varchar
DT - Datetime

The event can
occur prior to
the fatality, so
this is not
redundant

After running transform.py, both tables are fully populated in mysql:

```
mysql> select count(*) from fatalities;
+-----+
| count(*) |
+-----+
| 19288 |
+-----+
1 row in set (0.01 sec)

mysql> select count(*) from details;
+-----+
| count(*) |
+-----+
| 1682037 |
+-----+
1 row in set (0.64 sec)

mysql> select BEGIN_YEARMONTH,BEGIN_DAY,BEGIN_TIME,END_YEARMONTH,END_DAY,END_TIME, EPISODE_ID,EVENT_ID,STATE,STATE_FIPS,EVENT_TYPE,CZ_TYPE from details limit 5;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| BEGIN_YEARMONTH | BEGIN_DAY | BEGIN_TIME | END_YEARMONTH | END_DAY | END_TIME | EPISODE_ID | EVENT_ID | STATE | STATE_FIPS | EVENT_TYPE | CZ_TYPE |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 197009 | 4 | 1650 | 197009 | 4 | 1650 | NULL | 9881965 | ARKANSAS | 5 | Thunderstorm Wind | C |
| 197010 | 8 | 1655 | 197010 | 8 | 1655 | NULL | 10058129 | MISSOURI | 29 | Tornado | C |
| 197006 | 4 | 1615 | 197006 | 4 | 1615 | NULL | 10047491 | MISSISSIPPI | 28 | Thunderstorm Wind | C |
| 197006 | 11 | 2200 | 197006 | 11 | 2200 | NULL | 10056944 | MISSOURI | 29 | Thunderstorm Wind | C |
| 197007 | 1 | 2100 | 197007 | 1 | 2100 | NULL | 10052308 | MINNESOTA | 27 | Thunderstorm Wind | C |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)

mysql> select CZ_FIPS,CZ_NAME,WFO,CZ_TIMEZONE,INJURIES_DIRECT,INJURIES_INDIRECT,DEATHS_DIRECT,DEATHS_INDIRECT,DAMAGE_PROPERTY,DAMAGE_CROPS,SOURCE,MAGNITUDE,MAGNITUDE_TYPE,FLOOD_CAUSE,CATEGORY from details limit 5;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| CZ_FIPS | CZ_NAME | WFO | CZ_TIMEZONE | INJURIES_DIRECT | INJURIES_INDIRECT | DEATHS_DIRECT | DEATHS_INDIRECT | DAMAGE_PROPERTY | DAMAGE_CROPS | SOURCE | MAGNITUDE | MAGNITUDE_TYPE | FLOOD_CAUSE | CATEGORY |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 93 | MISSISSIPPI | NULL | CST | 0 | 0 | 0 | 0 | 0 | 0 | NULL | 0 | NULL | NULL | NULL |
| 197 | SCHUYLER | NULL | CST | 0 | 0 | 0 | 0 | 25K | 0 | NULL | 0 | NULL | NULL | NULL |
| 159 | WINSTON | NULL | CST | 0 | 0 | 0 | 0 | 0 | 0 | NULL | 0 | NULL | NULL | NULL |
| 9 | BARRY | NULL | CST | 0 | 0 | 0 | 0 | 0 | 0 | NULL | 0 | NULL | NULL | NULL |
| 93 | MEEKER | NULL | CST | 0 | 0 | 0 | 0 | 0 | 0 | NULL | 0 | NULL | NULL | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)

mysql> select BEGIN_RANGE,BEGIN_AZIMUTH,BEGIN_LOCATION,END_RANGE,END_AZIMUTH,END_LOCATION,BEGIN_LAT,BEGIN_LON,END_LAT,END_LON from details limit 5;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| BEGIN_RANGE | BEGIN_AZIMUTH | BEGIN_LOCATION | END_RANGE | END_AZIMUTH | END_LOCATION | BEGIN_LAT | BEGIN_LON | END_LAT | END_LON |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 | NULL | NULL | 0 | NULL | NULL | 36 | -90 | NULL | NULL |
| 0 | NULL | NULL | 0 | NULL | NULL | 40 | -93 | 40 | -93 |
| 0 | NULL | NULL | 0 | NULL | NULL | 33 | -89 | NULL | NULL |
| 0 | NULL | NULL | 0 | NULL | NULL | 37 | -94 | NULL | NULL |
| 0 | NULL | NULL | 0 | NULL | NULL | 45 | -94 | NULL | NULL |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
5 rows in set (0.00 sec)

mysql> select EVENT_NARRATIVE,EPISODE_NARRATIVE from details ORDER BY BEGIN_YEARMONTH desc limit 1;
+-----+-----+
| EVENT_NARRATIVE | EPISODE_NARRATIVE |
+-----+-----+
| In Danvers, water was up to the floor boards of a SUV (about 8 to 12 inches deep) on Village Post Road. | Strong southerly wind flow out ahead of an advancing cold front produced strong to damaging wind gusts and some heavy rain that caused some street and basement flooding. |
+-----+-----+
1 row in set (2.29 sec)

mysql>
```


Data is returning meaningful queries for analysis:

phpMyAdmin

Recent Favorites

New

Capstone

New

details

Columns

New

BEGIN_AZIMUTH

BEGIN_DAY

BEGIN_LAT

BEGIN_LOCATION

BEGIN_LON

BEGIN_RANGE

BEGIN_TIME

BEGIN_YEARMON

CATEGORY

CZ_FIPS

CZ_NAME

CZ_TIMEZONE

CZ_TYPE

DAMAGE_CROPS

DAMAGE_PROPER

DEATHS_DIRECT

DEATHS_INDIREC

END_AZIMUTH

END_DAY

END_LAT

END_LOCATION

END_LON

END_RANGE

END_TIME

END_YEARMONTH

EPISODE_ID

Server: localhost:3306 » Database: Capstone » Table: details

Browse Structure SQL Search Insert Export Import Privileges Operations Triggers

Show query box

⚠ Current selection does not contain a unique column. Grid edit, checkbox, Edit, Copy and Delete features are not available.

✔ Showing rows 0 - 24 (70 total, Query took 8.8693 seconds.)

```
SELECT EVENT_TYPE,SUM(DEATHS_DIRECT),SUM(DEATHS_INDIRECT),SUM(INJURIES_DIRECT),SUM(INJURIES_INDIRECT) FROM details GROUP BY 1 ORDER BY 2 DESC,3 DESC, 4 DESC, 5 DESC
```

1 > >> ☐ Show all Number of rows: 25 Filter rows:

+ Options

EVENT_TYPE	SUM(DEATHS_DIRECT)	SUM(DEATHS_INDIRECT)	SUM(INJURIES_DIRECT)	SUM(INJURIES_INDIRECT)
Tornado	3648	37	64662	268
Heat	2533	246	10293	210
Flash Flood	1695	60	6469	58
Rip Current	1140	8	813	2
Hurricane (Typhoon)	1097	25	1382	2409
Thunderstorm Wind	1018	87	11244	375
Excessive Heat	1005	152	5815	68
Lightning	887	47	5222	322
Flood	598	57	2387	51
Cold/Wind Chill	590	52	298	31
Avalanche	383	6	247	5
High Wind	350	93	1532	438
Wildfire	324	47	2171	454
High Surf	262	17	329	6
Strong Wind	243	64	551	74
Winter Storm	237	302	1444	1690
Extreme Cold/Wind Chill	206	34	26	1
Winter Weather	163	777	2368	3620
Heavy Snow	124	127	712	582
Debris Flow	121	3	242	2
Heavy Rain	119	109	279	355
Tropical Storm	109	87	391	20
Ice Storm	98	84	334	1487
Blizzard	97	90	411	141
Dense Fog	69	148	852	696