

Data Sleuth

Missing and Erroneous Data in International Trade Networks

P.G. Swartz

peter@panjiva.com

Panjiva, Inc.

www.panjiva.com

K.W. Soura

keith@panjiva.com

Overview

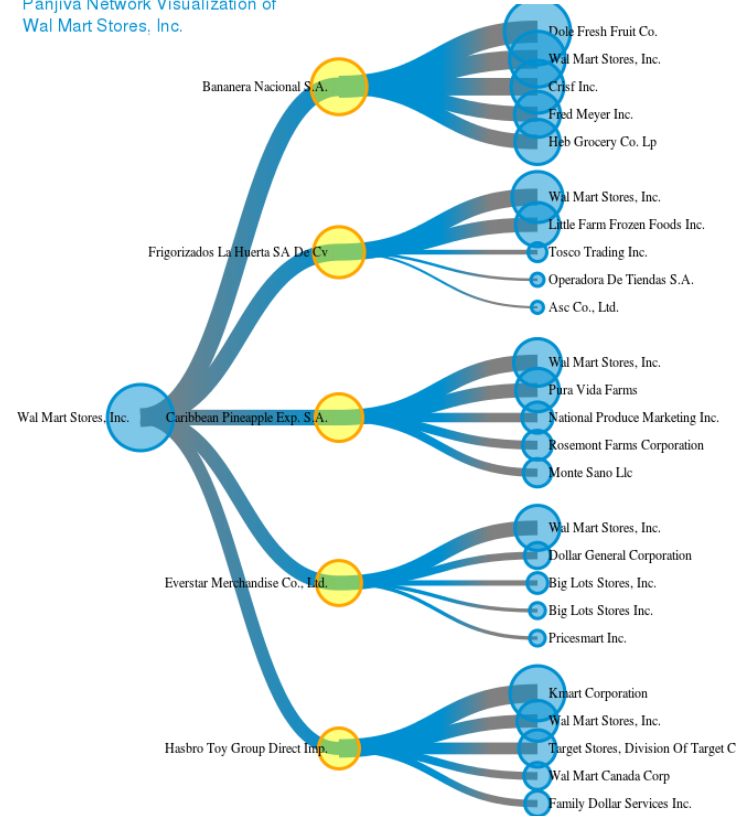
- Panjiva Introduction
- Importance of Missing / Incorrect Data
- Panjiva's Solutions to Missing Data:
 - Missing observations: Entire shipments
 - Missing values: Product category codes
 - Incorrect values: Weights

What is Panjiva?

Panjiva: Mapping Global Trade

- 13 Governments
 - US
 - China
 - Mexico
 - Brazil
 - + 9 South/Central America
- 190 countries
- 600 million records
- 7 million companies

Panjiva Network Visualization of Wal Mart Stores, Inc.



Transaction-Level Records

Contact Info

[hide](#)

Type	Name	Address
Consignee	SCHAEFFLER MEXICO S DE RL DE CV	RIO SAN LORENZO 888 CASTRO DEL RIO, (PARQUE TECNOINDUSTRIAL), IRAPUATO MX
Notify Party	Schaeffler Mexico S De Rl De Cv	RIO SAN LORENZO 888 CASTRO DEL RIO, (PARQUE TECNOINDUSTRIAL), IRAPUATO MX
Also Notify Party		
Shipper	KUNSTSTOFFTECHNIK ROS GMBH & INDUSTRIESTR. 14, KUEPS DE CO. KG	

Product Details

[hide](#)

Container	Commodity	Marks Description	HTS Codes
TTNU9816755	PLASTIC CAGE FOR BEARINGS BANQBRE2651356 HS-CODE 848299.	SCHAEFFLER MEXICO S DE RL DE CV IRAPUATO MEXIKO INV. 15002832, 15002833	HTS 8482.99

Transit Details

[hide](#)

Carrier Name	BLUE ANCHOR AMERICA LINE (BLUE ANCHOR LINE)
Carrier Details	Address: 10 EXCHANGE PL 19TH FLOOR, JERSEY CITY, New Jersey, United States, 07302 -- Phone: 2014135676
Ship Registered In	
Vessel Name	MAERSK KARLSKRONA
Voyage Number	533W
U.S. Port of Unlading	The Port of Charleston, Charleston, South Carolina
Foreign Port of Lading	Bremerhaven, Germany

Company Profiles From Transactions

★ FAVORITE

↓ GROUP

SHARE

Zhongde Equipment Co., Ltd.Shandong

ALL LOCATIONS - CHINA ▼ | SUPPLIER ✓

SummaryProductsShipmentsChinese ExportsCustomersCorporateContact

Products



See more product information

Shipments

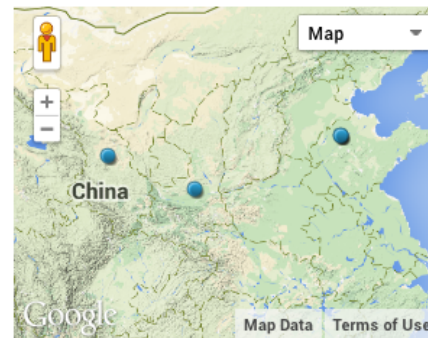
126

Shipments sent

TOP PRODUCT TERMS

beer	
brewing	
beer brewing	
brewing equipment	
beer brewing equipment	

COMPANY LOCATIONS

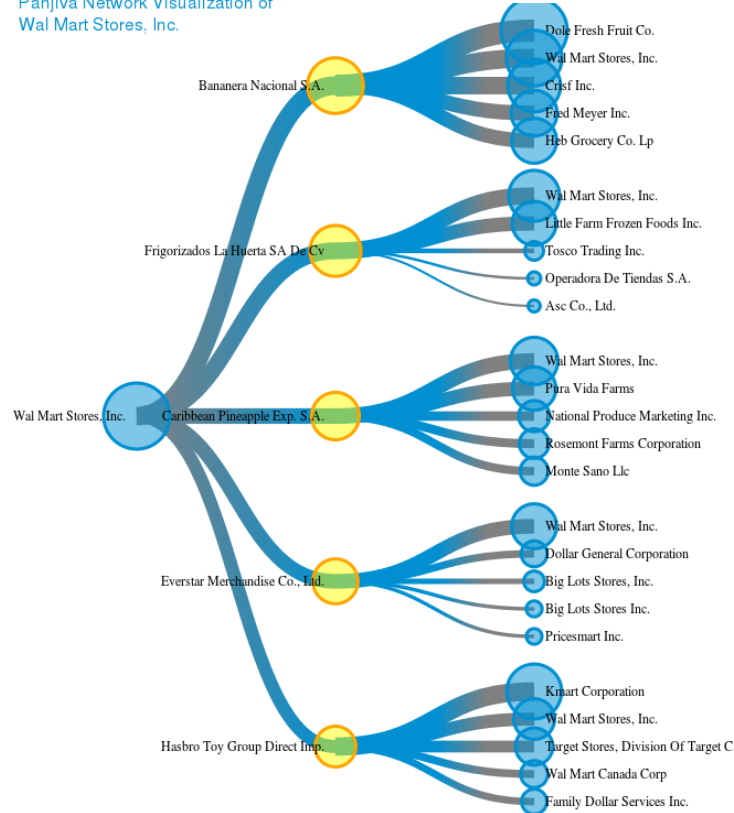


DATA SOURCES



Networks from Company Profiles

Panjiva Network Visualization of
Wal Mart Stores, Inc.



The Importance of Missing and Erroneous Data

Theoretical

- **Goal:** Real world statistical analysis
- **Obstacle:** Missing and erroneous data
- **Solutions:** Detect and mitigate
 - Missing observations
 - Missing values
 - Incorrect values

Missing Observations

- **Detect:**
 - Identify periods of anomalous record counts
- **Correct:**
 - Treat as missing values / erroneous values in aggregate analysis
- **Enable:**
 - Principled approach to abnormal periods

Detect Missing Observations

SHIPMENT RECORDS

698k ●

[hide details...](#)

KILOGRAMS

26B ●

[more details...](#)

VALUE OF GOODS

\$4.6B ●

[more details...](#)

Data source

Shipment Records

Total

698,853

100%



US Imports

578,480

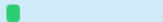
83%



Mexico Imports

60,093

9%



Chile Imports

56,482

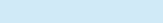
8%



Brazil Imports

2,429

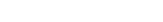
0%



China Imports

1,090

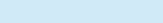
0%



Peru Imports

156

0%



Colombia Imports

55

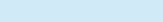
0%



Bolivia Imports

42

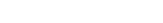
0%



Ecuador Imports

13

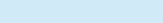
0%



Paraguay Imports

10

0%



Panama Imports

3

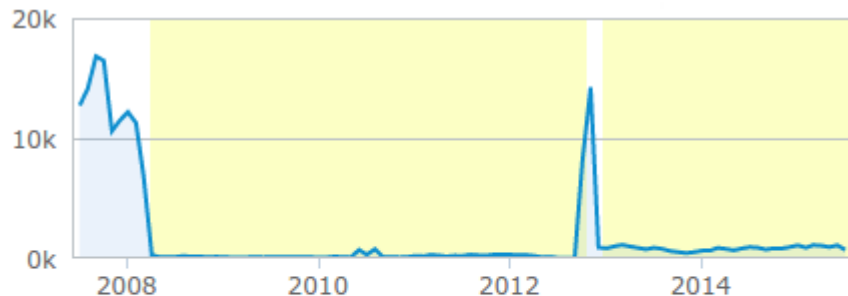
0%



Shipments

173,303

Shipments received



Wal Mart Stores, Inc. may have requested to hide their shipment activity from one or more data sources.

[Learn how to find hidden shipments](#)

Missing Values

- **Detect:**
 - Identify sparse fields
- **Correct:**
 - Impute these missing values
- **Enable:**
 - Unbiased and efficient analysis
 - No longer assume the values are Missing Completely at Random (MCAR)!
 - Query and investigate individual records

Fill in Missing Values

Shipments ▾ camisa ▾ 🔍

SHIPMENT RECORDS

883k ●

[hide details...](#)

KILOGRAMS

74B ●

[more details...](#)

VALUE OF GOODS

\$4.7B ●

[more details...](#)

Data source	Shipment Records		
Total	883,234	100%	<div></div>
Peru Imports	463,129	52%	<div></div>
Mexico Imports	338,250	38%	<div></div>
Chile Imports	37,587	4%	<div></div>
Uruguay Imports	21,557	2%	<div></div>
Panama Imports	11,429	1%	<div></div>
Bolivia Imports	3,890	0%	
Colombia Imports	3,339	0%	
Paraguay Imports	2,468	0%	
US Imports	1,308	0%	
Brazil Imports	277	0%	

Shipments ▾ shirt ▾ 🔍

SHIPMENT RECORDS

1.7M ●

[hide details...](#)

KILOGRAMS

100B ●

[more details...](#)

VALUE OF GOODS

\$9.5B ●

[more details...](#)

Data source	Shipment Records		
Total	1,669,338	100%	<div></div>
US Imports	957,771	57%	<div></div>
China Imports	294,321	18%	<div></div>
Peru Imports	229,842	14%	<div></div>
Mexico Imports	94,168	6%	<div></div>
Uruguay Imports	39,612	2%	<div></div>
Chile Imports	33,145	2%	<div></div>
Panama Imports	9,793	1%	<div></div>
Colombia Imports	4,300	0%	
Brazil Imports	3,970	0%	
Paraguay Imports	1,854	0%	
Bolivia Imports	525	0%	
Venezuela Imports	35	0%	
Ecuador Imports	2	0%	

Fill in Missing Values

Arrival Date	Consignee	Shipper	Weight	Container Info
				Commodity Description
15k KG				
2015-08-20	Gresham Imp. Gladstone Rd P O Box Cb 13931 Nassau Np Bahamas ☎ +1 242-424-2121 ☎ +1 242-396-1137 ☎ +1 242-461-0100	Car Junction Co., Ltd. Asaka, Japan ☎ +81 81484862301	15,920 KG (15.9 MT)	USED CARS - USED VEHICLESTOYOTA WISH ZNE10-03 09306 TOYOTA BREVIS JCG10-0069227 TOYOTA WIND OM MCV30-6040813 TOYOTA VITZ SCP90- 5035642
				Show 2 more containers...

Fill in Missing Values

Arrival Date	Consignee	Shipper	Weight	Container Info	
				Commodity Description	HTS Codes ⓘ
15k KG					
2015-08-20	Gresham Imp. Gladstone Rd P O Box Cb 13931 Nassau Np Bahamas ☎ +1 242-424-2121 ☎ +1 242-396-1137 ☎ +1 242-461-0100	Car Junction Co., Ltd. Asaka, Japan ☎ +81 81484862301	15,920 KG (15.9 MT)	USED CARS - USED VEHICLESTOYOTA WISH ZNE10-03 09306 TOYOTA BREVIS JCG10-0069227 TOYOTA WIND OM MCV30-6040813 TOYOTA VITZ SCP90-5035642	Classified 8703 Vehicles other than railway or tramway rolling stock, and parts and accessories thereof: Motor cars and other motor vehicles principally designed for the transport of persons (other than those of heading 8702), including station wagons and racing cars:
Show 2 more containers...					

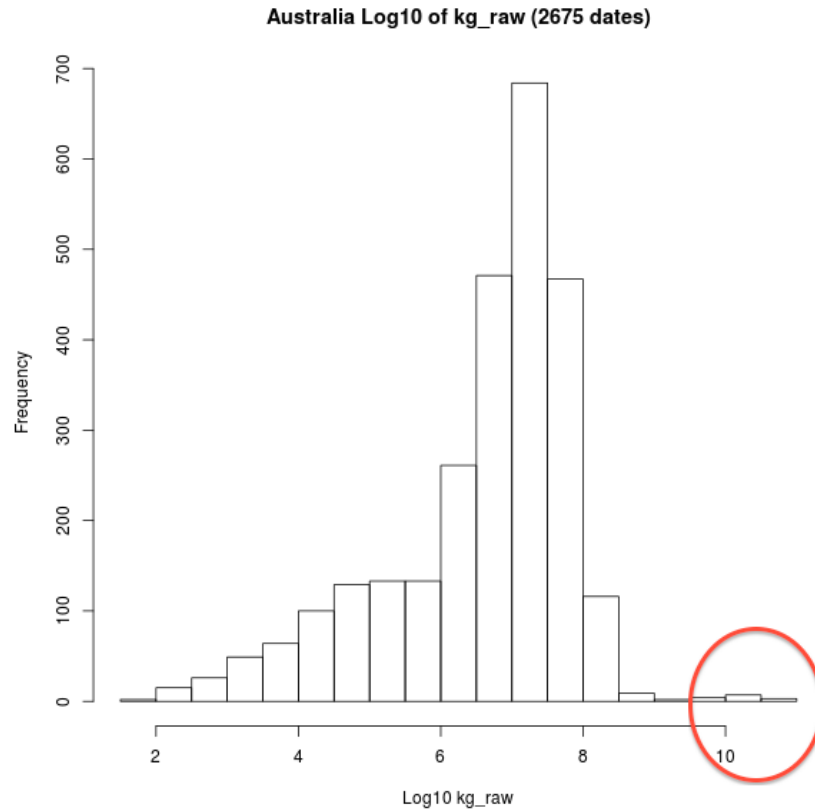
Erroneous Values

- **Detect:**
 - Identify abnormal values
- **Correct:**
 - Impute or domain specific fix
- **Enable:**
 - Analysis at the micro and macro level
 - Minimization of outliers

Correct Erroneous Values

Arrival Date	Consignee	Shipper	Weight	Container Info
				Commodity Description
			130M KG	
2015-07-14	Marathon Petroleum Co Findlay, Ohio, United States ☎ +1 713-629-6600 👤 ~770 Employee Contacts Available...	Saudi Arabian Oil Co. Dhahran, Saudi Arabia ☎ +966 3 872 0115 🌐 saudiamco.com	130,812,000 KG (130,812.0 MT) Weight greatly exceeded the capacity of a large bulk ship. Reduced by a factor of 1,000 from 130,812,000,000.	ARAB LIGHT CRUDE; API GRV 33.50; 961,471 BBLS

Correct Erroneous Values



Missing No More

Questions & Answers

- Contact us for more information
 - Peter G. Swartz, Data Scientist
 - peter@panjiva.com
 - Keith W. Soura, Technical Account Manager
 - keith@panjiva.com

Data Sleuth

Detecting and Recovering Missing Data in International Trade Networks

P.G. Swartz

peter@panjiva.com

Panjiva, Inc.

www.panjiva.com

K.W. Soura

keith@panjiva.com