

DATA QUALITY REPORT

Descriptive stats for continuous features - transposed table:

	count	mean	std	min	25%	50%	75%	max
MarketplaceId	10000.0	1.156462e+18	2.355318e+04	1.156462e+18	1.156462e+18	1.156462e+18	1.156462e+18	1.156462e+18
ProductId	10000.0	5.534192e+16	5.308389e+18	-9.172604e+18	-4.356483e+18	2.477570e+16	5.056101e+18	9.207147e+18
ListingPrice	10000.0	2.158817e+02	2.555817e+02	3.240000e+00	6.333000e+01	1.260200e+02	2.578925e+02	3.194320e+03
SellerFeedbackRating	10000.0	8.903970e+01	2.147030e+01	0.000000e+00	9.100000e+01	9.500000e+01	9.600000e+01	1.000000e+02
SellerFeedbackCount	10000.0	6.915795e+03	1.097018e+04	0.000000e+00	3.380000e+02	3.293000e+03	8.452000e+03	4.142000e+04
SellerId	10000.0	9.780680e+17	4.952648e+18	-9.136118e+18	-2.572278e+18	1.163115e+18	5.477945e+18	9.174048e+18
ShippingPrice	10000.0	1.239678e+01	2.636123e+01	0.000000e+00	0.000000e+00	7.500000e+00	1.364000e+01	7.052700e+02
ShippingTime_minHours	10000.0	5.713680e+01	8.245091e+01	0.000000e+00	2.400000e+01	2.400000e+01	9.600000e+01	6.720000e+02
ShippingTime_maxHours	10000.0	8.866320e+01	1.198274e+02	0.000000e+00	4.800000e+01	4.800000e+01	1.200000e+02	1.008000e+03

Descriptive stats for categorical features

	count	unique	top	freq
IsWinner	10000	2	0	9451
IsFeaturedMerchant	10000	2	1	8166
IsFulfilledByAmazon	10000	2	0	9632
ListingCurrency	10000	1	CAD	10000
ShippingCurrency	10000	1	CAD	10000
ShippingTime_availability	10000	1	NOW	10000
ShipsDomestically	10000	1	1	10000
ShipsFromCountry	6273	13	CA	3668
ShipsFromState	5880	24	ON	2220
SubCondition	10000	1	new	10000

Duplicated rows 114,
New shape (9886, 21)

Drop constant columns, denoted by 1 below:

```

Feature, UniqueValues
IsWinner, 2
MarketplaceId, 1
ProductId, 307
TimeOfOfferChange, 548
ConditionNotes, 436
IsFeaturedMerchant, 2
IsFulfilledByAmazon, 2
ListingPrice, 4976
ListingCurrency, 1
SellerFeedbackRating, 25
SellerFeedbackCount, 123
SellerId, 187
ShippingPrice, 1104
ShippingCurrency, 1
ShippingTime_minHours, 12
ShippingTime_maxHours, 11
ShippingTime_availability, 1
ShipsDomestically, 1
ShipsFromCountry, 14
ShipsFromState, 25
SubCondition, 1

```

New shape after dropping constant columns: (9886, 15)

Descriptive stats for continuous features - transposed table

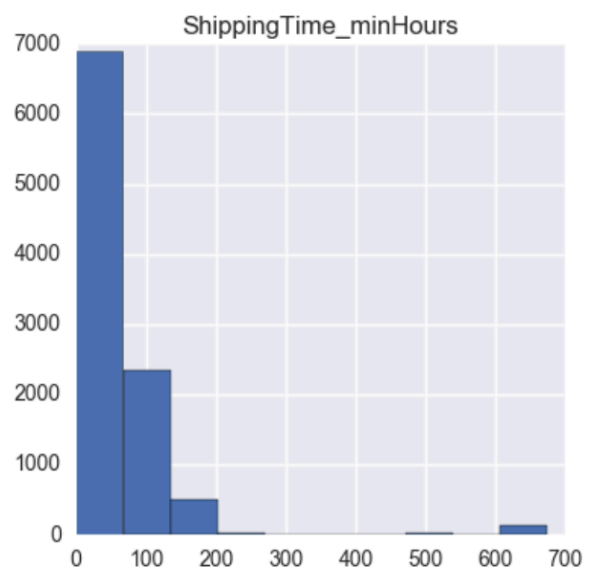
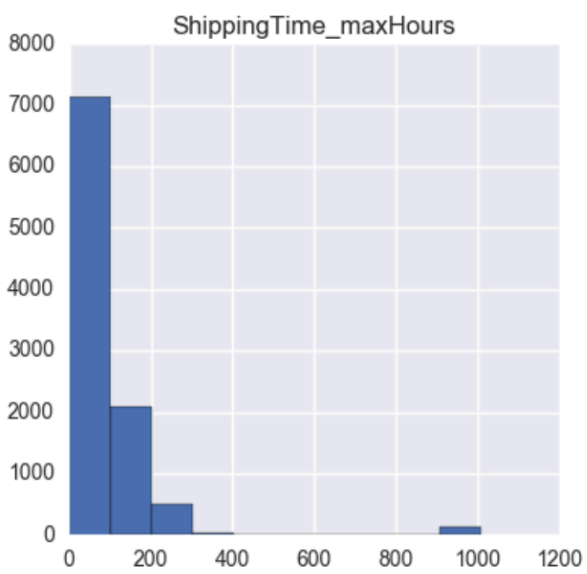
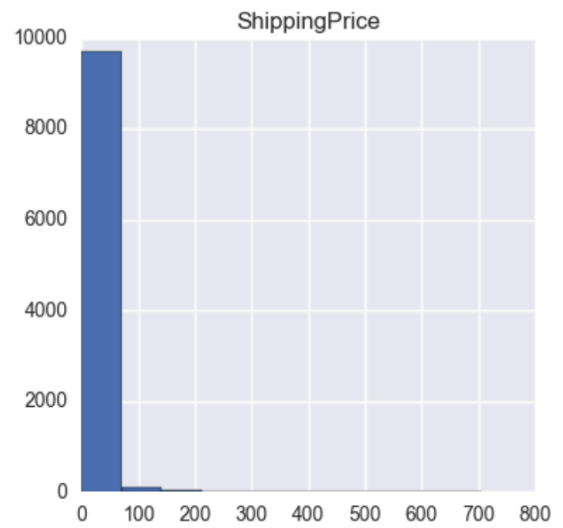
	count	mean	std	min	25%	50%	75%	max
ProductId	9886.0	6.104485e+16	5.315584e+18	-9.172604e+18	-4.356483e+18	2.477570e+16	5.056101e+18	9.207147e+18
ListingPrice	9886.0	2.164802e+02	2.565794e+02	3.240000e+00	6.363000e+01	1.259900e+02	2.578625e+02	3.194320e+03
SellerFeedbackRating	9886.0	8.897593e+01	2.155962e+01	0.000000e+00	9.100000e+01	9.500000e+01	9.600000e+01	1.000000e+02
SellerFeedbackCount	9886.0	6.910402e+03	1.091824e+04	0.000000e+00	3.380000e+02	3.293000e+03	8.452000e+03	4.142000e+04
SellerId	9886.0	9.641003e+17	4.919222e+18	-9.136118e+18	-2.572278e+18	1.163115e+18	5.477945e+18	9.174048e+18
ShippingPrice	9886.0	1.243470e+01	2.647657e+01	0.000000e+00	0.000000e+00	7.500000e+00	1.368000e+01	7.052700e+02
ShippingTime_minHours	9886.0	5.726644e+01	8.280107e+01	0.000000e+00	2.400000e+01	2.400000e+01	9.600000e+01	6.720000e+02
ShippingTime_maxHours	9886.0	8.887477e+01	1.203972e+02	0.000000e+00	4.800000e+01	4.800000e+01	1.200000e+02	1.008000e+03

Descriptive stats for categorical features

	count	unique	top	freq
IsWinner	9886	2	0	9339
IsFeaturedMerchant	9886	2	1	8090
IsFulfilledByAmazon	9886	2	0	9519
ShipsFromCountry	6217	13	CA	3655
ShipsFromState	5851	24	ON	2211

Plot a histogram of the continuous features





Bar plots for all the categorical features

