


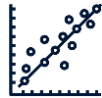
# What Is the Price of Data? A Measurement Study of Commercial Data Marketplaces

 180 surveyed DMs, 213,964 products from 2,015 DPs > 10,772 paid data products > 4,162 products with prices.

Marketplace	#Data Products	#Paid products	#Sellers
Advaneo	198,743	1	N/A
AWS	4,013	2,515	262
DataRade	1,592	1,592	1,262
Knoema	158	158	142
DAWEX	160	160	79
Carto	8,182	5,283	42
Crunchbase	16	14	15
Veracity	115	95	38
Refinitiv	214	185	76
Other data providers	771	769	29



We built a cross-DM database and trained classifiers to compare between DMs



We tested 9 regressors and optimized 4 of them to get at least one  $R^2 > 0.78$  for predicting the price of data product of different categories of data

Table 2: Score of data product classifiers

	Accuracy	Precision	Recall	$F_1$ Score
Test - Financial	0.93	0.97	0.81	0.88
Test - Retail	0.95	0.96	0.88	0.91
Val. - Financial	0.89	0.72	0.88	0.79
Val. - Retail	0.78	0.81	0.68	0.74

Model	Financial			Marketing			Healthcare			All		
	$R^2$	MAE	MSE	$R^2$	MAE	MSE	$R^2$	MAE	MSE	$R^2$	MAE	MSE
RF	0.85	0.2	0.14	0.86	0.21	0.13	0.78	0.25	0.15	0.84	0.23	0.16
kN	0.78	0.31	0.26	0.74	0.33	0.24	0.77	0.26	0.17	0.69	0.37	0.31
GB	0.82	0.23	0.16	0.8	0.28	0.19	0.73	0.27	0.19	0.79	0.3	0.22
DNN	0.73	0.33	0.35	0.77	0.30	0.22	0.68	0.26	0.18			



30% of paid products cover the US, and prices range up to USD\$ x00M.

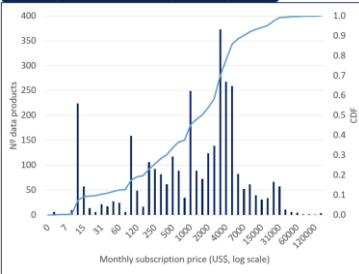


Telecom, Manufacturing and Automotive median prices above the global (x2.6, x2.3 and x2) in AWS. 3% monthly growth in #products

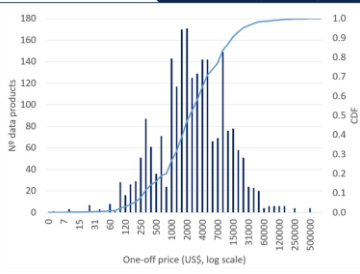


It is mostly 'what' (description and categories) and 'how much' (volume & update rate) data is being traded that determine its price

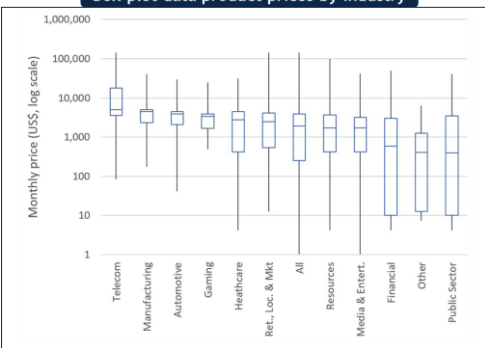
Subscription-based data product prices



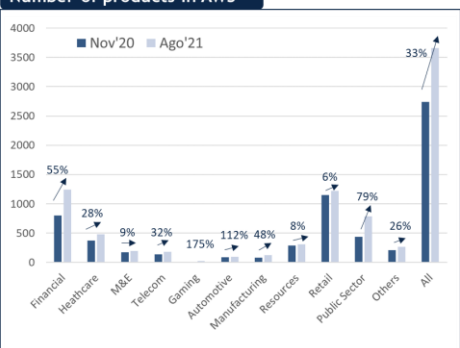
One-off data product prices



Box-plot data product prices by industry



Number of products in AWS



Predicting power by feature group

