

SALE PRICE PREDICTIONS

CRYSTAL BALL CREATED BY

MICHAEL POWERS

DATA SAMPLE

Item ID	Price Purchased	Price Sold	Quantity for Stock	Date Purchased by Store	Date put on Shelf	Date Purchased by Customer	Store Name	Weekly Sales	Monthly Sales	Quarterly Sales	Yearly Sales	Expiration Date	Seasonal?	Initial Release Date	Brand	Subcategory	Allergy Warning	Damaged?	On Sale
6731809754	\$3.79	\$5.94	442	4/3/2020	4/4/2020	4/6/2020		\$4,324.80	\$18,435.38	\$61,820.04	\$229,184.87	4/3/2022	NO	1/20/1962			YES	NO	NO
4218221611	\$6.74	\$8.10	250	10/22/2019	10/23/2019	10/28/2019		\$2,263.66	\$7,368.30	\$21,322.42	\$76,004.50	10/21/2021	NO	1/13/2000			NO	NO	NO
9635054334	\$3.16	\$2.57	177	2/10/2020	2/11/2020	2/16/2020		\$631.17	\$2,529.44	\$7,802.24	\$35,806.17	2/9/2022	YES	5/13/1966			NO	NO	YES
2210022356	\$2.16	\$2.52	175	3/12/2018	3/13/2018	3/13/2018		\$624.36	\$2,982.45	\$9,462.60	\$32,419.74	3/11/2020	NO	4/9/1962			NO	NO	NO
9128591178	\$6.10	\$8.15	495	11/10/2017	11/11/2017	11/15/2017		\$4,435.49	\$20,994.26	\$69,897.37	\$250,220.98	11/10/2019	NO	10/17/1965			YES	NO	NO
7541551294	\$4.09	\$5.29	246	12/27/2019	12/28/2019	12/30/2019		\$1,942.01	\$9,098.35	\$25,967.83	\$101,544.45	12/26/2021	NO	3/16/1987			YES	NO	NO
4282330370	\$3.28	\$5.14	224	4/16/2021	4/17/2021	4/18/2021		\$1,325.50	\$6,159.81	\$18,281.35	\$60,485.12	4/16/2023	NO	2/17/1993			NO	NO	NO

FEATURES OF THE DATA

- Item ID
- Price Purchased
- Quantity in Stock
- Date Purchased by Store
- Date put on Shelf
- Date Purchased by Customer
- Weekly Sales
- Monthly Sales
- Quarterly Sales
- Yearly Sales
- Expiration Date
- Is Product Seasonal
- Initial Release Date
- Allergy Warning
- Is Product Damaged
- Is Product on Sale

WHAT MODELS WERE USED?

- Dummy (dr)
- Linear (lr)
- Random Forest (rfr)
- Decision Tree (dtr)
- Light Gradient Boost Model (lgbm)
- Cat Boost (cbr)

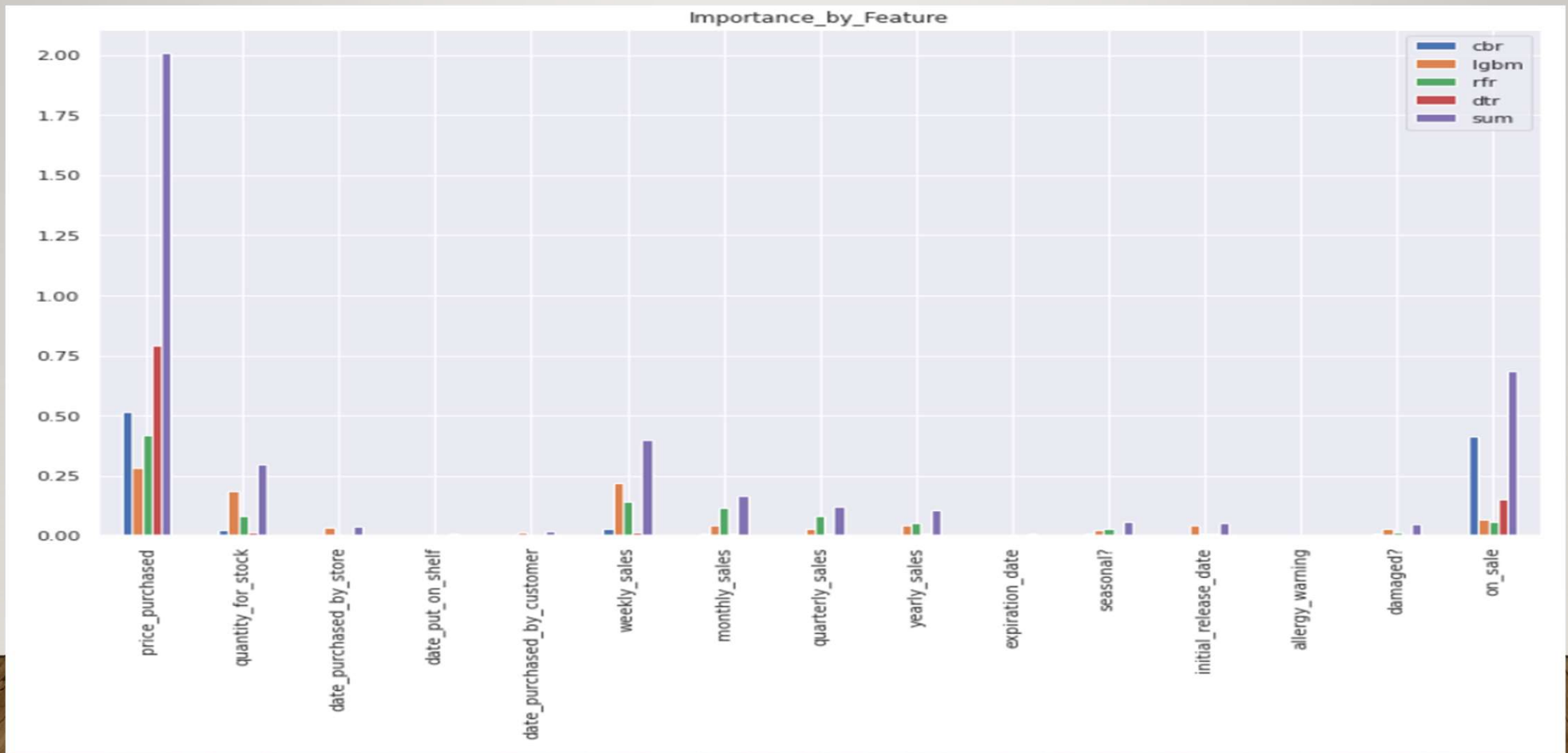
PRICE SOLD VERSUS MODELS' PREDICTIONS

Item ID	price_sold	dr_predictions	lr_predictions	dtr_predictions	rfr_predictions	cbr_predictions	lgbm_predictions
1000653731.00	\$4.86	\$4.69	\$5.99	\$5.21	\$5.95	\$5.99	\$5.99
1000967045.00	\$4.23	\$4.69	\$4.26	\$4.20	\$4.34	\$4.28	\$4.28
1001151067.00	\$5.24	\$4.69	\$6.09	\$4.92	\$5.24	\$5.40	\$5.40
1009039405.00	\$3.13	\$4.69	\$3.49	\$3.79	\$3.44	\$3.34	\$3.34
1009650054.00	\$6.14	\$4.69	\$6.52	\$6.39	\$6.74	\$6.64	\$6.64
1011172386.00	\$4.67	\$4.69	\$4.45	\$5.28	\$4.37	\$4.36	\$4.36
1013195956.00	\$3.63	\$4.69	\$4.27	\$4.48	\$3.92	\$4.01	\$4.01
1013798447.00	\$6.55	\$4.69	\$7.74	\$7.70	\$7.73	\$7.59	\$7.59

HOW IMPORTANT ARE THE FEATURES?

	cbr	lgbm	rfr	dtr	sum
price_purchased	0.51378	0.28242	0.41947	0.79089	2.00655
on_sale	0.41220	0.06468	0.05599	0.15465	0.68753
weekly_sales	0.02433	0.22203	0.14040	0.01262	0.39938
quantity_for_stock	0.02172	0.18760	0.07751	0.01103	0.29786
monthly_sales	0.00901	0.03911	0.11150	0.00557	0.16519
quarterly_sales	0.00431	0.02817	0.07916	0.00514	0.11679
yearly_sales	0.00337	0.04163	0.05197	0.00525	0.10221
seasonal?	0.00532	0.02046	0.02779	0.00168	0.05526
initial_release_date	0.00019	0.04063	0.00563	0.00555	0.05199
damaged?	0.00544	0.02799	0.01403	0.00031	0.04776
date_purchased_by_store	0.00009	0.03140	0.00388	0.00177	0.03715
date_purchased_by_customer	0.00008	0.01105	0.00396	0.00158	0.01667
expiration_date	0.00008	0.00000	0.00389	0.00168	0.00565
date_put_on_shelf	0.00007	0.00000	0.00395	0.00155	0.00556
allergy_warning	0.00001	0.00285	0.00085	0.00073	0.00444

HOW IMPORTANT ARE THE FEATURES?



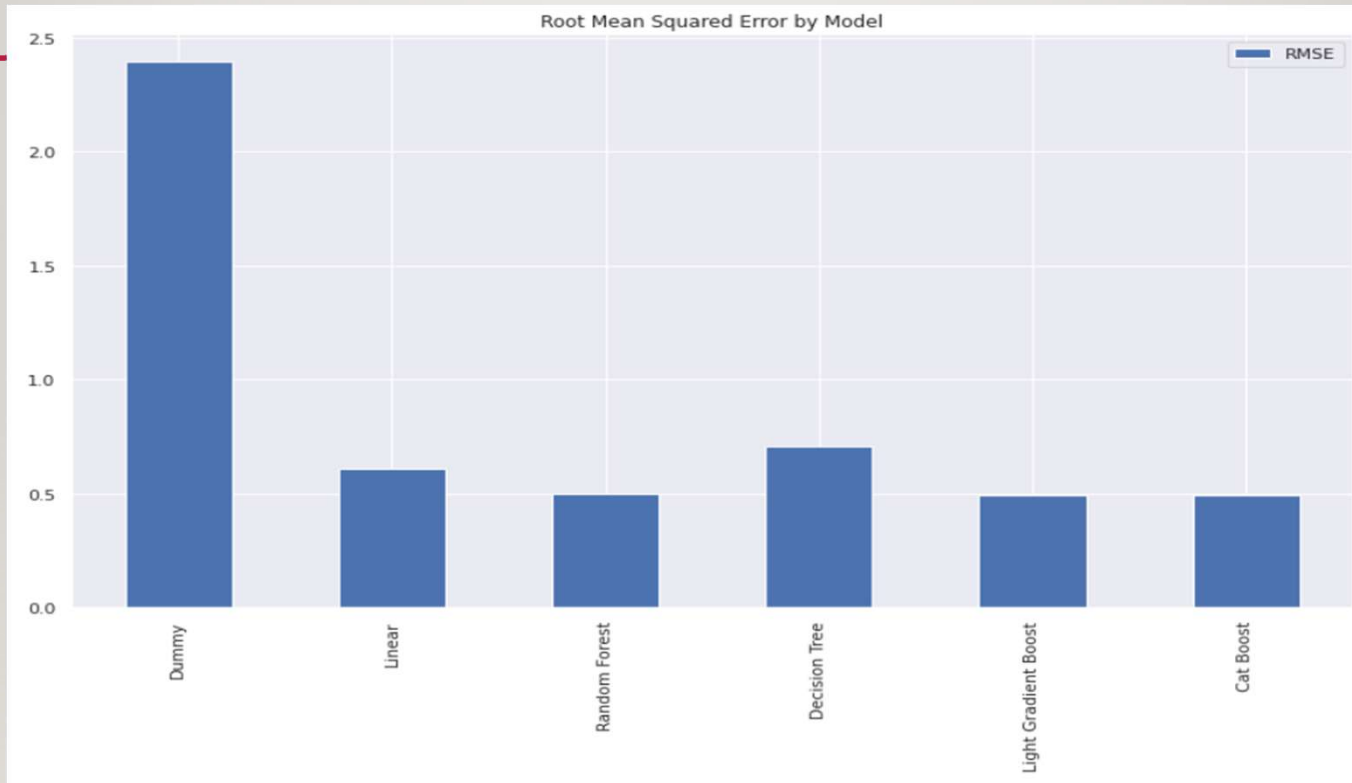
AVERAGE ERROR AND TRAINING TIME PER MODEL

	RMSE	Fit Time
Dummy	2.393998	0.00
Linear	0.610317	0.01
Random Forest	0.500453	17.96
Decision Tree	0.704177	0.91
Light Gradient Boost	0.492373	18.49
Cat Boost	0.492373	12.71

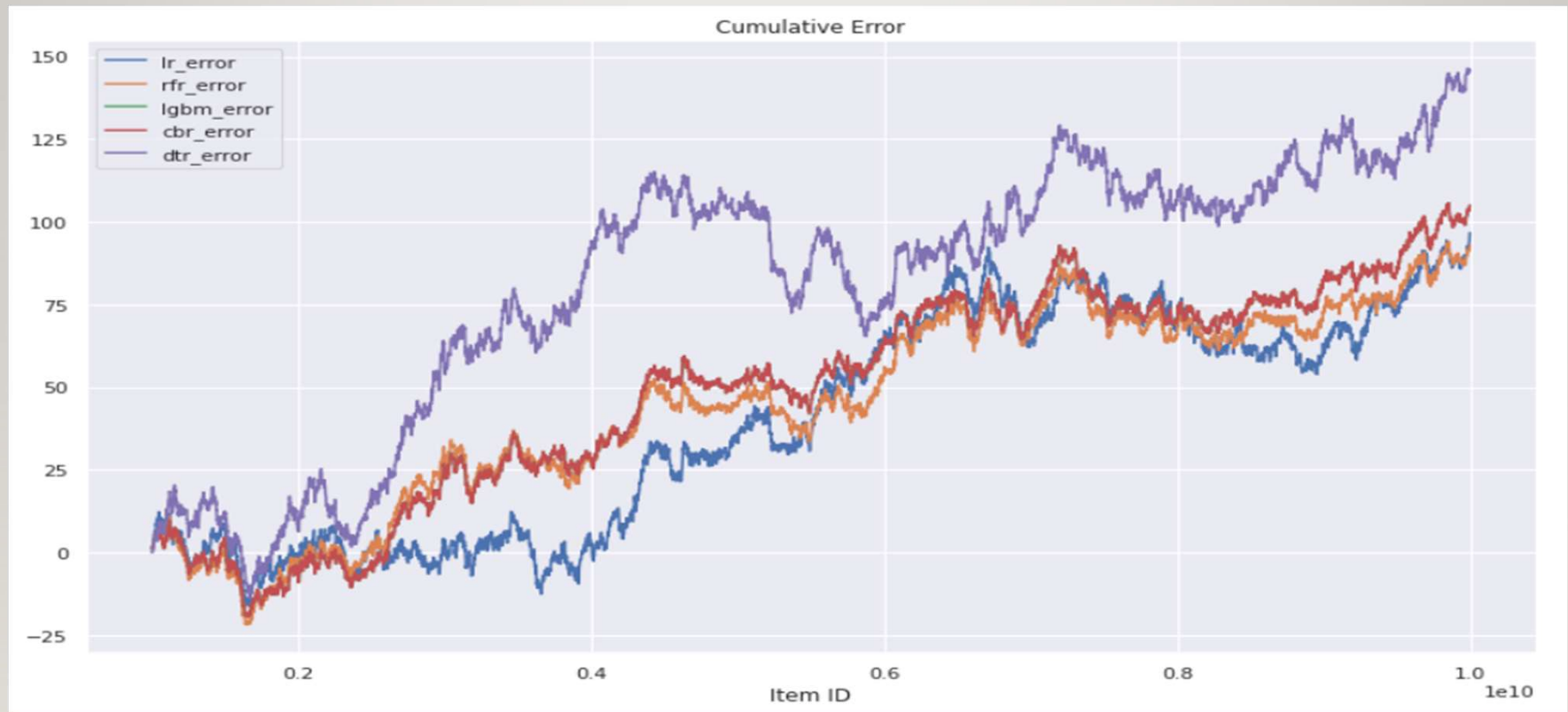
RMSE: Root Mean Squared Error

Fit Time: Time in seconds it takes
to train each model

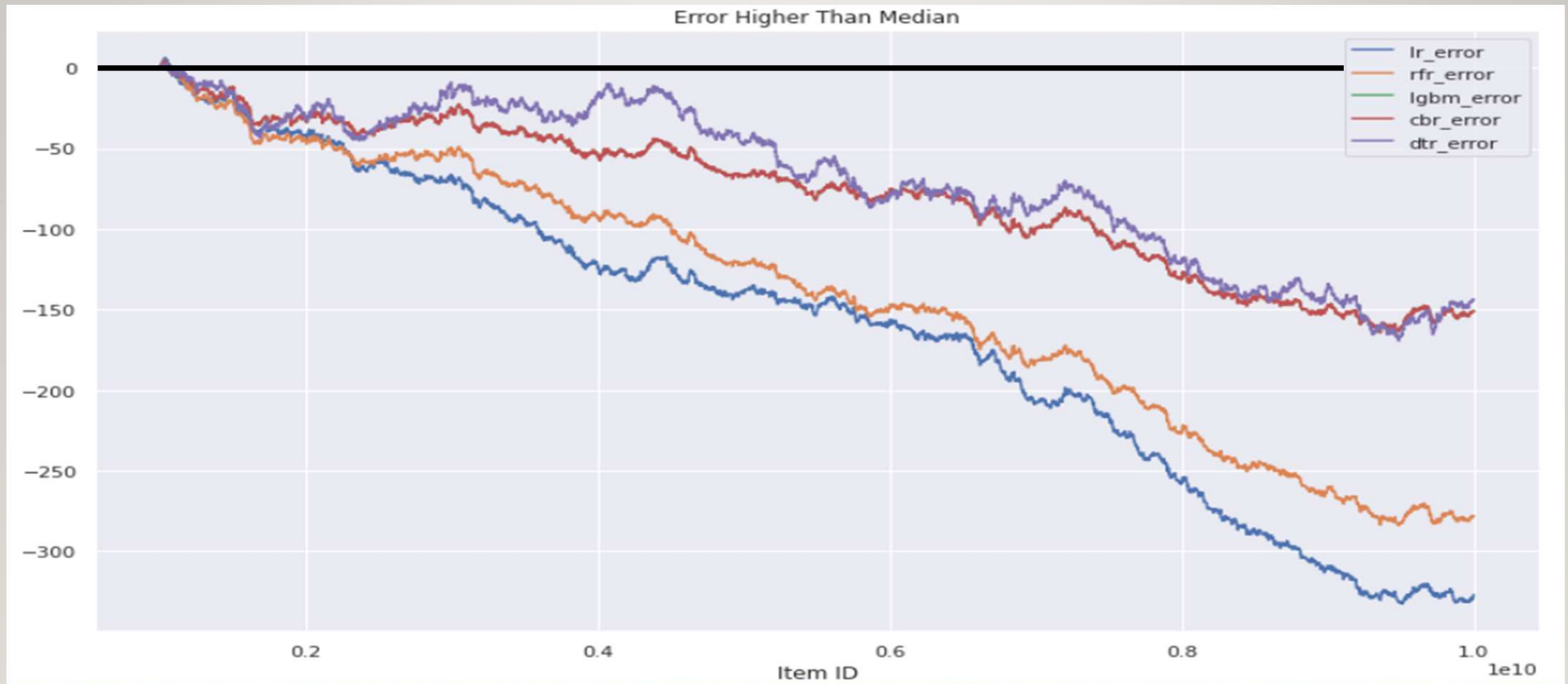
AVERAGE ERROR GRAPHICS



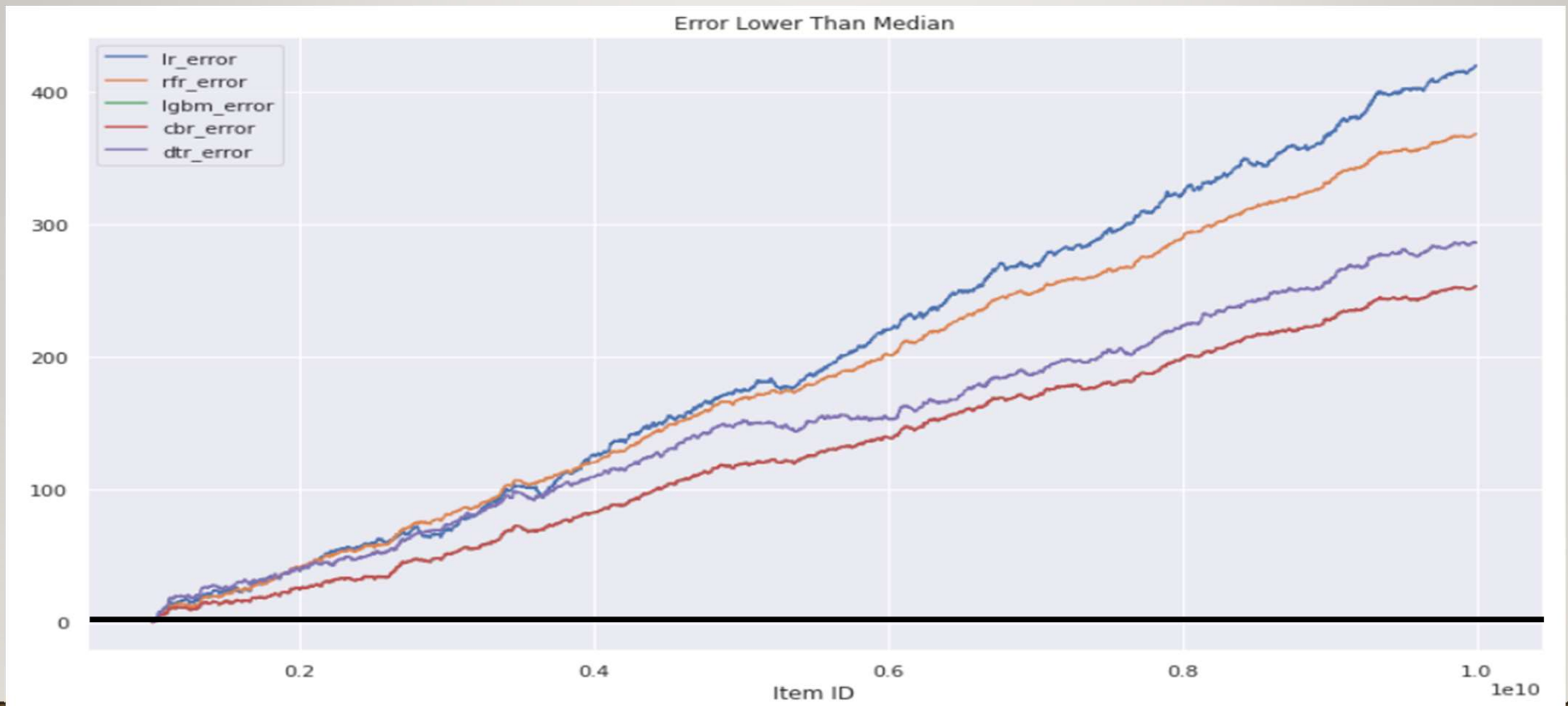
CUMULATIVE ERROR PER MODEL OVER FULL DATASET



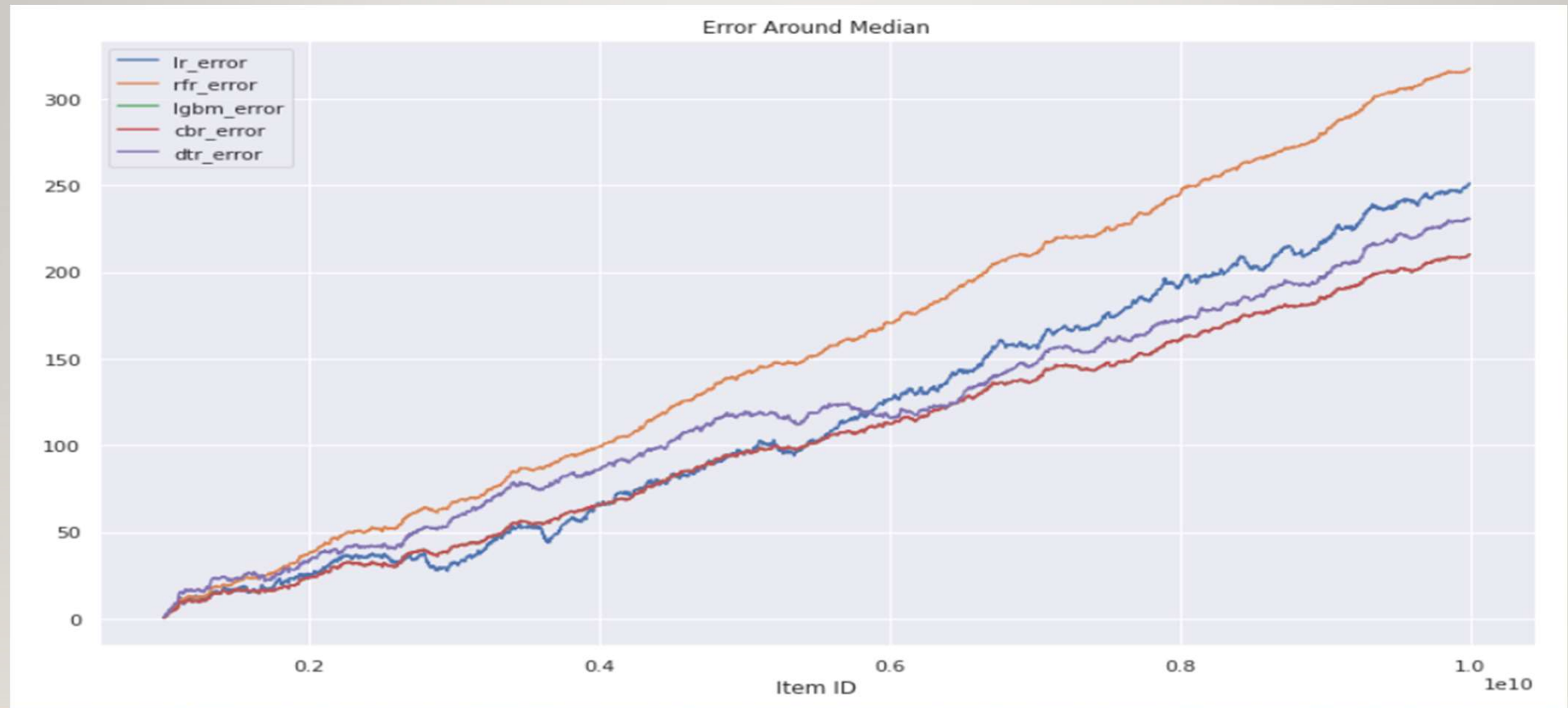
CUMULATIVE ERROR OVER HIGH PRICED ITEMS



CUMULATIVE ERROR OVER LOW PRICED ITEMS



CUMULATIVE ERROR OVER AVERAGE PRICES



FINAL THOUGHTS

Item ID	price_sold	dr_predictions	lr_predictions	dtr_predictions	rfr_predictions	cbr_predictions	lgbm_predictions
1000653731.00	\$4.86	\$4.69	\$5.99	\$5.21	\$5.95	\$5.99	\$5.99
1000967045.00	\$4.23	\$4.69	\$4.26	\$4.20	\$4.34	\$4.28	\$4.28
1001151067.00	\$5.24	\$4.69	\$6.09	\$4.92	\$5.24	\$5.40	\$5.40
1009039405.00	\$3.13	\$4.69	\$3.49	\$3.79	\$3.44	\$3.34	\$3.34
1009650054.00	\$6.14	\$4.69	\$6.52	\$6.39	\$6.74	\$6.64	\$6.64
1011172386.00	\$4.67	\$4.69	\$4.45	\$5.28	\$4.37	\$4.36	\$4.36
1013195956.00	\$3.63	\$4.69	\$4.27	\$4.48	\$3.92	\$4.01	\$4.01
1013798447.00	\$6.55	\$4.69	\$7.74	\$7.70	\$7.73	\$7.59	\$7.59

Links to Full Data

<https://docs.google.com/spreadsheets/d/1cDBpQyWlqcsF8IyabU4Hh2Wsj5QapHXI5QAvAfFwYII/edit?usp=sharing>

<https://docs.google.com/spreadsheets/d/1VKdWZUh7Wr05njyzJTlpuckOTPYUrrRiQi5AqRzj4HU/edit?usp=sharing>

<https://github.com/michaelpowers8/Several-Grocery-Items.git>