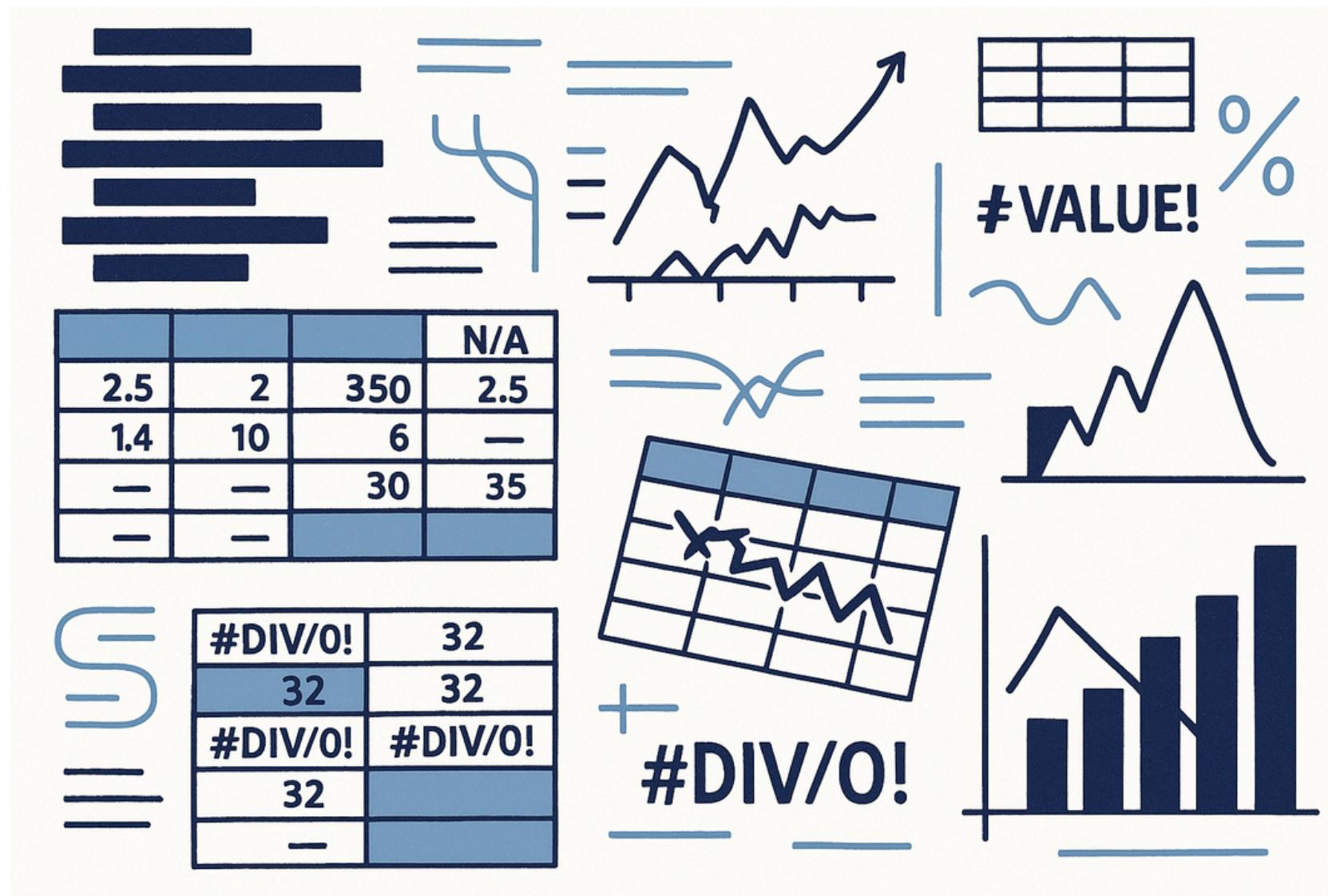


Arch's Mission. My Contribution.

Harsh Sharma | Github | harshnarayan.sharma@duke.edu

The Risk of Good Enough Data



- The world of private investing is full of data but it is often unusable.
- Inconsistent formats, missing details, and manual errors prevent a clear view of performance.
- It creates a liability for clients and a bottleneck for the business.

A Closer Look at the Risk

DATE	ID	REGION	CATEGORY	VALUE
05/14/203	1032	West	—	\$500
06/22/203	529	—	Office	32000
N/A	267	South	Other	\$450
05/21/203	336			\$700
05/13/203	466			\$1000
04/202	223			\$100
12/14/203	248			\$200
02/28/203	269			\$500
06/15/203	397	Solar	Avanage	\$500
05/18/203	169			\$200
10/16/203	833			00
06/23/203	941			00
ABC	DRT			00

Data Gaps

Key columns are riddled with missing values, leaving crucial information up to chance

Inconsistent Values

The VALUE column is a mix of dollars and thousands, making financial analysis impossible

Unstructured Chaos

Related data is scattered across multiple files, forcing a manual, time-consuming join

The Automated Path to Data Integrity



My role is to be the Data Quality Associate who builds a foundation of trust. I developed an automated pipeline to transform this raw data into a reliable, single source of truth.

My Solution: The Python Pipeline

Building on a Solid Foundation

BEFORE

DATE	ID	REGION	CATEGORY
05/14/203	1032	West	—
06/22/203	529	—	Office
N/A	267	South	Other
05/21/203	336	F"II	\$700
05/13/203	466	9100	\$1000
12/14/203	248	Abc	\$200
02/21/203	369	—	\$500
02/28/203	397	Solar	\$500
06/15/203	169	Avange	\$500
10/16/203	833	DRT	\$200
06/23/203	941	—	\$200
ABC	DRT	—	

AFTER

DATE	ID	REGION	VALUE	
05/14/203	1021	—	\$500	
06/22/203	1022	Offiee	\$32,000	
6/5/2031	1023	—		Financial Accuracy All VALUE metrics are normalized, enabling confident, reliable analysis
6/14/2031	1023	Stos		
6/21/203	1024	Olat		
6/22/203	1025	Solar		
6/23/203	1026	Avange		Complete Data Critical gaps are filled providing a complete picture of holdings and reports
6/24/203	1027	Stoal		
6/25/203	1028	Solar		
6/26/203	1029	CrIL		
6/27/203	1035			
633/203	1035			Actionable Insight The data is unified into a single, clean file, allowing teams to focus on strategy, not manual work

Why I'd Love to be Part of the Mission



I absolutely love the idea behind creating this platform of becoming the one stop destination in the alternative investments landscape by through technology and collaboration.

My project was a small taste of my passion for the kind of work that makes that mission possible.

Thank you! I would love the opportunity to contribute to the future you're building.

THANK YOU

Harsh Sharma | Github | harshnarayan.sharma@duke.edu