Field performance analysis by Ehsan Ghani:

Motivation:

This document demonstrates an analytical framework for achieving that by simulating one month of field performance for 50 sales executives. The goal is to provide data-driven evidence to support the team's development, offering an objective outlook on how to get the best out of each team member and reducing the bias of purely intuitive conclusions.

The methodology is designed to be highly versatile. While this analysis covers 50 executives over one month, the model is effective across any team size or timeframe. It can also be attuned to the specific KPIs and metrics that your company values most once I become more familiar with them.

It's important to preface that this report is a demonstration of my statistical modelling process, conducted prior to a deep understanding of a specific business' model. My aim is to support the team with actionable, numerical evidence, not to produce confusing models. I am aware this is a preliminary insight into the business and am eager to learn and adapt this analysis to ensure nothing is overlooked.

Process of data simulation, choices in data, benefit of randomness in data and proposed method of data collection:

When starting this task, it was essential to simulate a dataset that accurately reflects real-world field performance. The following parameters were defined based on industry research and conversations with partners:

. **Doors Knocked/Houses Reached:** This metric monitors the activity level of each sales executive. I set a monthly range between 900 and 1,600 and applied a normal distribution. This implies a realistic model where most salespeople perform near the average, with fewer outliers at the high and low ends of activity.

. **Contact Rate (Doors Answered):** This represents the percentage of knocks that result in a door being answered. This was set as a random probability for each executive, ranging from 30% to 55%, dependent on their Doors Knocked. This range assumes the team is operating in a territory with a reasonable chance of finding residents at home. Naturally, in real data this wouldn't be a necessary factor, but to account for realism in the simulation I deemed it necessary.

. **Sales Conversion:** This is the percentage of contacts (answered doors) that result in a sale. For this simulation, the sales conversion probability for each executive was set to a random value between 5% and 15%.

. **Average Deal Size:** To ground the model in a familiar context, the deal size was based on previous work with EE broadband services. The total value of a 24-month broadband contract (with monthly costs between £28.99-£38.99 and higher as contract continues) can result in an average deal size in the range of £1,007-£1,205, accounting for various high-end packages or setup fees.

. **Total Revenue:** This is derived directly from the number of sales and the randomly distributed deal size, reflecting the varied performance across a real team.

The Role of Randomness and Model Versatility

Randomness was used to simulate the natural variance in performance found in any sales team. This approach creates a realistic dataset with a wide range of outcomes, ensuring that the resulting analysis is robust and not based on uniform or perfect data. While any single simulation is just one of many possibilities, the analytical methods applied are universally applicable.

Proposed Method of Data Collection

The proposed method for data collection is designed to be low-friction and not disrupt the sales team's daily operations. My suggestion would be to integrate simple logging into the current CRM or use a tool like Google Forms or Jotform to automatically populate a central spreadsheet for analysis.

Important remark:

The initial findings from this analysis were triangulated with machine learning models to ensure their validity. For example, the four performance quadrants were initially defined by team averages, but a K-Means clustering algorithm later confirmed this structure by independently identifying four distinct performance groups that closely aligned with those quadrants.

Data insight:

Upon simulating the data, I imported to excel (the data could be in any format, but I went with the typical facet), processed and cleansed via Python, here's a small sample:

	Sales_Executives	Doors_Knocked(Month)	Contacts_Lead	Sales	Total_Revenue	Contact_Rate	Sales_Conversion	Performance_Tier	Activity_Level
0	EXEC001	1338	549	71	76262.907469	41.03	12.93	Top Performer	Medium Activity
1	EXEC002	1238	483	58	64208.442534	39.01	12.01	High Performer	Medium Activity
2	EXEC003	1258	667	67	72830.897570	53.02	10.04	High Performer	Medium Activity
3	EXEC004	1438	489	39	40486.866747	34.01	7.98	Needs Improvement	High Activity
4	EXEC005	1395	544	65	67610.918614	39.00	11.95	High Performer	High Activity

The performance tier is cut across 5 tiers , with placement determined solely by each executive's total revenue contribution. Because revenue is a direct derivative of sales, I used it as the primary performance indicator to avoid multicollinearity/complications in the model.

Next tier was activity level split from low-medium-high activity, determined by the contact factors.

The tiers are relative, meaning they are determined by an executive's rank compared to their peers, not by fixed performance targets. This allows to reflect performance irrespective of general business. In the event of an underwhelming month of business, the model still reflects the true performance of each executive.

With these measures of raw effort and objective performance, I can further begin to understand what drives an executive's unique performance on a case-by-case basis and what could drive their ability across the board.

Top performers and what drives their status:

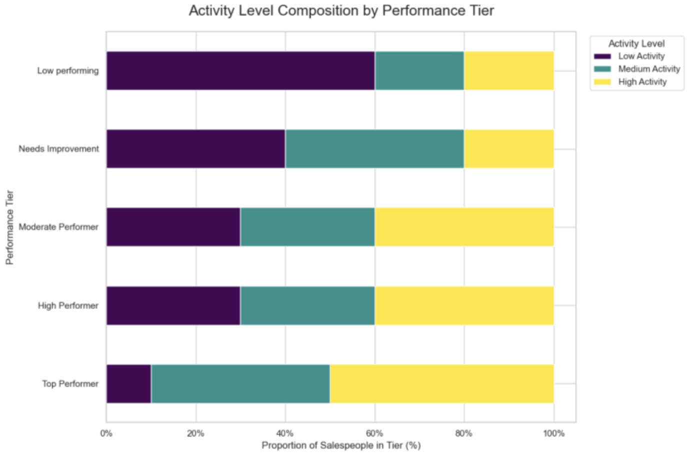
```
--- Top 5 Performers by Overall Rank ---
                  Revenue_Rank  Conversion_Rank  Activity_Rank  Overall_Rank
Sales_Executives
EXEC036                    2.0             12.0            1.0      5.000000
EXEC049                    1.0              2.0           18.0      7.000000
EXEC013                    3.0             19.0            2.0      8.000000
EXEC014                    5.0              9.0           11.0      8.333333
EXEC021                   14.0              7.5            5.0      8.833333
```

When creating the ranking system, the rank was determined on:
. Cumulative revenue contribution i.e. 'what they add to the pie'
. average sales conversion over the month recorded
. Activity levels recorded by their contact factors
This system rewards strengths inclusive of revenue and what could drive them further.

The current data suggests that amongst the top 5 there are clear hard workers/volume players and adept converters; executives 36 and 13 being dependent on their high activity and sheer volume of sales to boost their revenue contributions, whereas executive 49 is extremely adept at closing sales, but low activity is a hinderance.

These are two different conversations between the same category of 'top performers'. One may suggest that by executive 49 moderately increasing their activity, they could shorten the gap of first placement, as to not place a hinderance on their conversion and therefore revenue. For executives 36 and 13, further coaching on closing technique on sales could also contribute to that balance. The margins between executives we'd would wish to see is that of executives 14 and 21, small gaps but clear indications of areas for improvement.

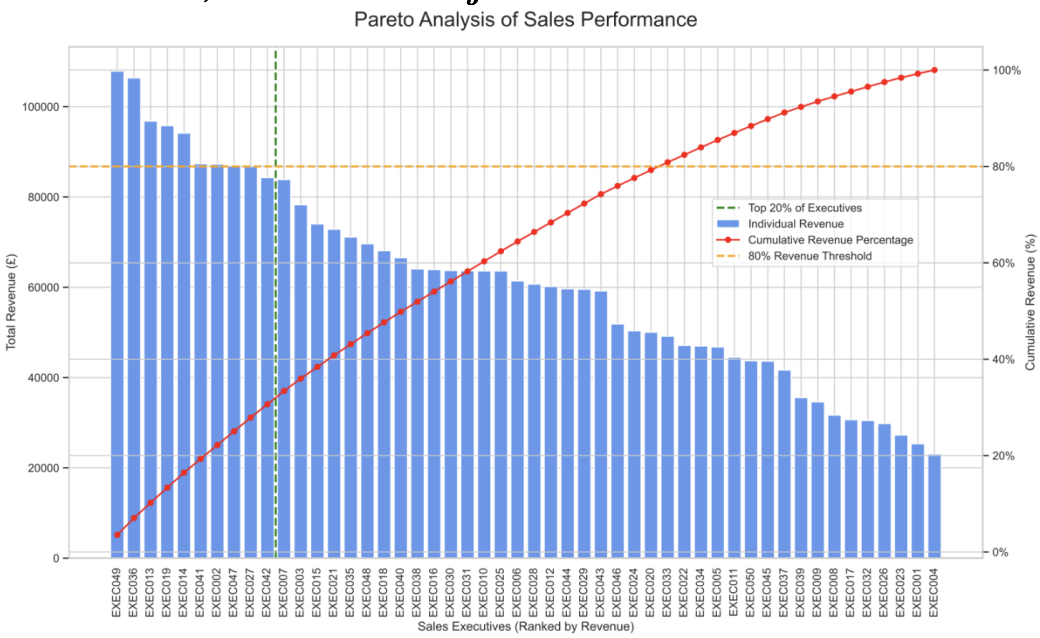
Where performance and activity levels lie in proportion:



Performance_Tier	Activity_Level	
Low performing	Low Activity	60.00%
	Medium Activity	20.00%
	High Activity	20.00%
Needs Improvement	Low Activity	40.00%
	Medium Activity	40.00%
	High Activity	20.00%
Moderate Performer	High Activity	40.00%
	Low Activity	30.00%
	Medium Activity	30.00%
High Performer	High Activity	40.00%
	Low Activity	30.00%
	Medium Activity	30.00%
Top Performer	High Activity	50.00%
	Medium Activity	40.00%
	Low Activity	10.00%

The data reveals a natural positive correlation between a salesperson's activity level and their performance tier. This relationship is most evident at the extremes: 90% of Top Performers are in the medium or high activity groups, establishing high effort as a near-essential prerequisite for elite success. Conversely, 60% of low performers are in the low activity category, this could be the consequence of the induction of trainees. If necessary, specific executives are extractable from the data to know certainly if they are trainees or established workers. Crucially, the identical activity profiles of moderate and high performers show that once a solid baseline of effort is established, skill (specifically conversion ability) becomes the key differentiator for advancing through the ranks. The level as to how strong this correlation is will be determined by the correlation heatmap found later.

Pareto chart, the indication of balance across revenue:

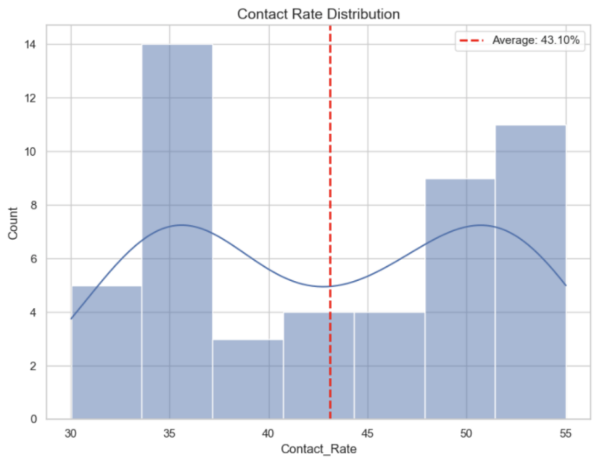


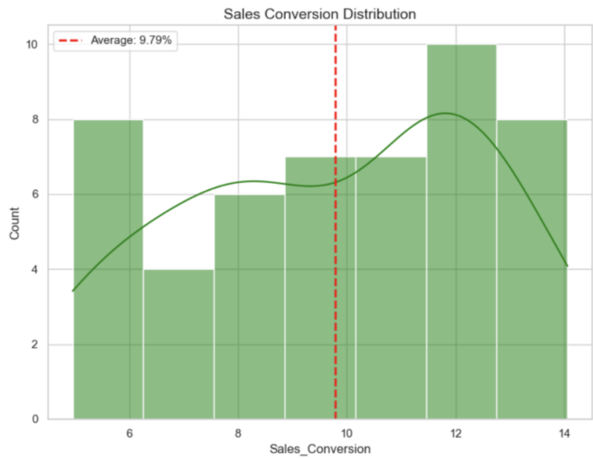
The chart above represents the cumulative contribution towards revenue. From the graph, it is shown that the top 34 (68%) of executives contribute to 80% of the revenue. This highlights the dependency on these performers staying consistent with their activity, conversions and contribution to revenue. Intuitively, we wish to restructure the balance as to increase the baseline performance across the team, so there's less risk and dependency on top performers. If say 45 of the 50 executives were to contribute to 80% of the revenue, this is sub-optimal as the other 5 (most likely top performers) would have the burden of being over 2x productive to contribute 20% of the entire revenue. Hence why the goal is to increase performance all-round to lessen the dependency on the minority, not just increase the number of executives contributing to the 80% . I can continuously monitor the balance with updated data.

Performance quadrant:

Here is a visual breakdown of performance with respect to sales conversion and efficiency of activity (contact rate). We can see with reflect to placement and size of the bubbles the varying performance across the team, in that it is fragmented across efficiency and activity. From the bottom left, it's clear the executives need some sort of foundational training, so that by the next collection of data we can begin to see anywhere from marginal to large increases in either conversion or activity. The bottom right and top left exhibit opposite strengths, high conversion and high activity respectively. Training these two groups with respect to the factors that limit them would be beneficial than addressing the training as a whole. The top right naturally are the optimal performers, it could be perhaps to maintain their ability to convert or reach clients effectively, they must continue in the manner they have, as to potentially mess with their methodology may come at the cost of their strength. Furthermore, just as the pareto analysis conveyed, restructuring the balance across performance would take the weight of top performers and moderately reduce the risk within the business.

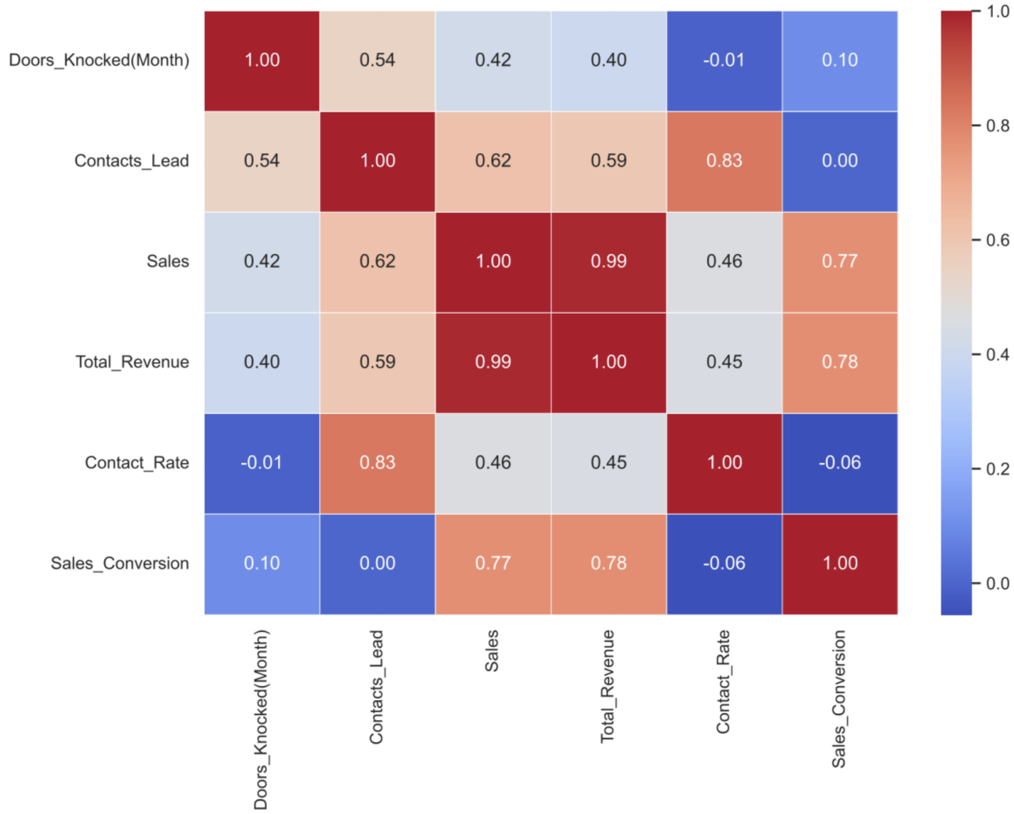
Distribution of sales conversion and contact rate:





I thought it was necessary to include the distributions of factors above to also explain how the averages produced in businesses often lead to misinterpretations of performance. From the graphs we can see that often two extremes of the data converge to an average that may tell a better story than the one that lies beneath the surfance. For example, when defining a sales conversion range from 5-15%, the average defined converged to the midpoint of approximately 10% and yet, the crucial insight is that this average is misleading because very few salespeople actually perform at this level. In this instance, it would be optimal to focus on segmented training by the groups shown in the performance quadrant, to make performance more predictable and less dependable on top conversion performers. Once again, dismissing wholesale training and segmenting across varying groups in strength of performers.

Correlation of metrics:



As I've found earlier, there was a positive correlation between activity and performance, the extent of how strong this correlation is in relation to other factors is found above. There is a 0.4 correlation between Doors knocked (activity) and total revenue (performance), this is significantly weaker than the behavioural correlation of sales conversion and total revenue (0.78). What this offers in terms of context is the confirmation that sales conversion is a more reliable indicator for top performance, a small margin increase in sales technique where necessary would go much further than simply increasing the level of activity.

This isn't to completely denote the importance of increasing an individual's activity, the benefit of the system is that it highlights an executive's profile across 3 factors, they may very well need to increase their activity to highlight their innate ability to close sales, and therefore have this reflect in their contribution towards revenue. Interestingly, the executive with the #1 rank in Sales Conversion did not make it into the top 5 by overall rank, suggesting that sales skill alone, without sufficient activity, is not enough to be a top overall contributor.

Overview of this data:

. **Implement a segmented coaching Program:** Dismiss wholesale training and instead use the performance quadrant to create targeted coaching plans.

. **Prioritize closing skills in profile-training:** Focus development resources on improving sales conversion, as it is the most reliable indicator of top performance. Whilst there is a behavioural correlation between revenue and activity, **efficiency in performance is preferred over an increase in activity.**

. **Real data and time to learn = better insights for your company:** As I continue to learn what a company values and work with real data, the analytical models can be refined to provide even more targeted and actionable recommendations for the team.