Feature response relationship



We are working on Moody's dataset with multiple features and responses for each market. It is quarterly data that means, that every quarter, new record come for each market. It is real estate dataset.

I am currently working with feature name average_hourly_earning_private and response vacancy. Help me understand how it works

Excellent — let's unpack this carefully and deeply.

You're working with Moody's real estate dataset (quarterly, multi-market, multi-feature), and focusing on the relationship between:

- Feature: average_hourly_earning_private
- **Response:** vacancy (vacancy rate of properties)



Step 1. Understanding the Feature

average_hourly_earning_private

This represents the average hourly earnings of private sector employees — i.e., what workers in the private sector earn per hour on average. It reflects:

- - Wage growth
 - Labor market strength
 - Inflationary pressures
 - Consumer spending potential

In Real Estate Context

Private sector earnings drive **demand-side fundamentals** of property markets:

Economic Effect	Mechanism	Impact on Vacancy
Higher wages → higher disposable income	People can afford rent or mortgage more easily.	↓ Vacancy (less empty space)
Higher wages → business cost pressure	Firms might reduce hiring or delay expansion if wages rise too fast.	↑ Vacancy (especially in offices)
Higher wages → inflationary environment	Central banks might raise rates \rightarrow higher borrowing costs \rightarrow less construction or slower investment.	Mixed effect depending on market phase
Sector-specific effect	Certain markets (tech, finance hubs) are more sensitive to wage trends.	Varies by market





🔃 Step 2. How It Interacts with Vacancy

Vacancy = unoccupied real estate space proportion (e.g., % of units or sqft available but not leased).

The relationship is often **non-linear** and **lagged**, because labor and property markets adjust slowly.

Typical Dynamics:

1. Lagged Response:

Wage growth now may influence vacancy 1-3 quarters later — as tenants renew leases, consumers adjust spending, and developers react.

• Use lag features: avg_hourly_earning_private_lag1, lag2, etc.

2. Differential Impact by Property Type:

- Office: correlated with corporate profitability and employment → more sensitive
- Retail: follows consumer spending
- **Apartment:** follows household income → negative correlation with vacancy
- Industrial: follows production & e-commerce → less directly tied to wages

3. Cross-market sensitivity:

Markets with high-income workers (NY, SF) may react differently than lower-wage markets.



Step 3. What to Expect Empirically

If you run a regression or correlation:

Property Type	Expected Sign of Coefficient (approx.)	Interpretation
Apartment	Negative	Higher wages $ ightarrow$ more affordability $ ightarrow$ lower vacancy
Retail	Negative or Mixed	Higher wages \rightarrow more spending, but inflation can offset
Office	Negative (with lag)	Higher earnings → strong firms → more space demand
Industrial	Weak/Indirect	Driven more by logistics than wages



🧩 Step 4. Suggested Feature Engineering

To make this feature richer and more "industry-grade", you can derive advanced variables:

Category	Engineered Feature	Economic Meaning
Growth rate	<pre>qoq_change = (current - previous)/previous</pre>	Detect wage momentum
Lag features	lag_1, lag_2, lag_4	Capture delayed vacancy response
Deviation from trend	actual - rolling_mean(4q)	Wage shocks (unexpected change)

Category	Engineered Feature	Economic Meaning
Ratio features	<pre>avg_hourly_earning_private / avg_hourly_earning_public</pre>	Wage gap between private and public sectors
Elasticity indicator	<pre>Interaction: avg_hourly_earning_private * employment_rate</pre>	Labor income power jointly
Real wage	avg_hourly_earning_private / CPI	Adjust for inflation
Market relative feature	(market_value - national_avg)	Local wage competitiveness
Volatility	std(earnings over last 4 quarters)	Stability of income in a region

Step 5. How to Analyze

1. Correlation & Lags:

- Compute correlation of vacancy with lagged earnings features.
- Plot cross-correlation function (CCF) to find optimal lag.

2. Regression Example:

$$Vacancy_t = \beta_0 + \beta_1 \cdot AvgHourlyEarningPrivate_{t-1} + \epsilon_t$$

or multivariate:

 $Vacancy_t = f(AvgHourlyEarningPrivate_{t-1}, Employment_t, CPI_t, InterestRate_t, ...)$

3. Causal Interpretation:

- Test if wage shocks cause vacancy changes using Granger causality or causal discovery.
- Be careful with endogeneity wage growth may itself depend on local economic health (simultaneous causality).

Step 6. Real-World Example

In **2021–2022**, U.S. private hourly earnings rose sharply post-COVID:

- Residential rents **increased** → lower vacancy in apartments.
- Office sector saw **mixed results** wage growth didn't offset remote work.
- Retail vacancy **improved** in suburban areas due to higher household income.

So the relationship is **contextual**, but the **sign and timing** depend on market and property type.

Would you like me to go one step further and show:

- 🔢 a **feature engineering pipeline** (in code or pseudo-code), or
- 👔 a **visual causal interpretation** diagram (how earnings connect to vacancy through mediators like employment, inflation, and interest rates)?



Help with out of box features engineering for this column.

Perfect — let's go for "out-of-the-box" feature engineering ideas for average hourly earning private, the kind of advanced transformations that top real estate analytics teams (like Moody's, CBRE, or CoStar) use to extract deeper insights from labor market variables.

We'll move beyond just lags and growth rates — and create macro-financial, behavioral, and cross-market features that connect private earnings to vacancy in real estate markets.

1. Momentum & Trend Features

Feature	Formula / Idea	Real Estate Insight
Earnings Growth (QoQ & YoY)	(E_t - E_{t-1}) / E_{t-1} and (E_t - E_{t-4}) / E_{t-4}	Measures wage momentum — rapid wage growth may precede tightening labor markets $\rightarrow \downarrow$ vacancy
Earnings Acceleration	(QoQ_growth_t - QoQ_growth_{t- 1})	Detects turning points — sudden acceleration signals strong demand
Rolling Average & Trend Deviation	E_t - rolling_mean(E, 4)	Captures "shock" in wages beyond trend

2. Lagged & Lead Features (Behavioral Delay)

Feature	Idea	Use
Lag_1, Lag_2, Lag_3, Lag_4	Earnings from previous quarters	Vacancy responds to wage shifts with delay
Forward Earnings (Lead)	Use future data in modeling scenario testing	Simulates impact of projected wage growth

3. Real Income–Adjusted Features

Feature	Formula	Why It Matters
Real Average Hourly Earnings	E_t / CPI_t	Adjusts for inflation — nominal wage gains may not translate into higher purchasing power
Disposable Income Proxy	<pre>E_t * EmploymentRate_t * (1 - TaxRate_t)</pre>	Links earnings to actual spending capacity
Affordability Index (Earnings-to-Rent)	E_t / MedianRent_t	Measures how affordable housing is for workers → strong negative link with vacancy

💤 4. Cross-Market and Spatial Features

Feature	Formula / Idea	Real Estate Logic
Market Relative Wage Index	(E_market_t / E_national_t)	Measures if a market's wage level is above or below national average
Regional Spillover Wage Index	Average wages of neighboring markets (weighted by distance)	Captures cross-market labor competition
Rank-based Feature	Market's percentile rank in wage growth	Helps detect "hot labor markets" driving local demand

5. Structural / Sectoral Interaction Features

Feature	Combination	Interpretation
Wage × Employment	E_t * EmploymentPrivate_t	Total wage bill \rightarrow better proxy for aggregate income
Wage × Education Employment	<pre>E_t * Education_Health_Services_Employment</pre>	High-earning, stable sector → sustained demand
Wage × Office Employment Share	<pre>E_t * (OfficeEmployment / TotalEmployment)</pre>	Captures wage pressure specifically from office-using sectors
Wage × Consumer Confidence Index	Interaction with sentiment	Measures how wage perception amplifies spending behavior



The Second Seco

Feature	Description	Insight
Hodrick-Prescott (HP) Filter Trend Component	Extract smooth trend from wage series	Identifies long-term wage trajectory vs short-term noise
HP Cycle Component	Short-term cyclical deviations	Detects cyclical overheating or softening of labor market
Wage Volatility (Rolling Std)	<pre>std(E_t over past 4 or 8 quarters)</pre>	Wage instability can reduce confidence → higher vacancy risk



7. Macro-Financial Derived Features

Feature	Formula / Idea	Economic Rationale
Wage–Interest Rate Spread	E_t / FedFundsRate_t	How fast wages are rising vs borrowing cost
Wage–Productivity Gap	E_t / IndustrialProductionIndex_t	Indicates wage pressure beyond productivity — can squeeze firms
Wage–Inflation Gap	E_t / CPI_t_growth	High real wage gains = more purchasing power → ↓ vacancy
Wage Elasticity Proxy	ΔVacancy / ΔEarnings (estimated empirically)	Quantifies market sensitivity to wage shifts



8. Demand & Affordability Hybrid Features

Feature	Description	Why Useful
Housing Demand Index	<pre>(Real_Wage * PopulationGrowth) / HousingSupplyGrowth</pre>	Integrates wage growth into demand fundamentals
Office Absorption Proxy	<pre>(E_t * OfficeEmployment_t) / OfficeStock_t</pre>	Wage growth + job creation \rightarrow office space absorption
Retail Sales Potential	<pre>(E_t * Employment_t * ConsumerConfidence_t)</pre>	Combines wage strength with sentiment and jobs

b 9. Nonlinear Transformations

Transformation	Reason	
log(E_t)	Stabilizes variance, interpretable in % changes	
sqrt(E_t) or E_t^2	Capture diminishing or amplifying returns	
z-score normalized by market	Compare across cities/markets consistently	
interaction with time trend	To detect structural breaks (e.g., post-COVID wage regime shift)	

10. AI/ML Derived Features

Technique	Description	Usage
Principal Component on Labor Features	Combine wages, employment, productivity into latent "Labor Market Strength"	Captures multi-dimensional effect
K-Means on Wage Patterns	Cluster markets by wage trajectory	Different cluster behaviors can explain vacancy variation
Residual from Wage Forecast Model	actual - predicted	Wage surprises or shocks as driver of short-term vacancy changes

🔇 Practical Example (Pseudo-Code Style)

```
python
# Assume df has columns: ['market', 'quarter', 'avg_hourly_earning_private', 'CPI',
'employment', 'median_rent']
df['wage_qoq'] = df.groupby('market')['avg_hourly_earning_private'].pct_change(1)
df['wage_yoy'] = df.groupby('market')['avg_hourly_earning_private'].pct_change(4)
df['wage_accel'] = df['wage_qoq'] - df.groupby('market')['wage_qoq'].shift(1)
df['real_wage'] = df['avg_hourly_earning_private'] / df['CPI']
df['affordability_index'] = df['avg_hourly_earning_private'] / df['median_rent']
df['wage_volatility'] = df.groupby('market')
['avg_hourly_earning_private'].rolling(4).std().reset_index(level=0, drop=True)
df['market_relative_wage'] = df['avg_hourly_earning_private'] / df.groupby('quarter')
['avg_hourly_earning_private'].transform('mean')
```

Would you like me to now suggest which of these engineered features typically have the strongest **predictive relationship with vacancy** (based on evidence from commercial real estate models)? That way, I can help you prioritize the top 5-7 most powerful transformations before you start coding them.



Sorry I was asking can you look for some more top notch features engineering which top real estate industry use

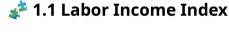
Excellent clarification — you're asking for **next-level**, "top-tier" feature engineering — the kind that's used by elite real estate analytics teams (think Moody's Analytics, CoStar, BlackRock, JLL Research, and REIS).

Let's go well beyond standard econometric features and get into the macro-structural, behavioral, and financial-engineering-inspired transformations that these firms actually use in their proprietary vacancy and rent models.

We'll keep the focus on average_hourly_earning_private as the base feature and design industry-grade engineered variables that capture how wage dynamics ripple through labor, consumption, credit, and space demand.



📔 1. Labor Market Power Features (Macro Strength Indicators)



Formula:

Labor Income Index_t = Avg Hourly Earnings_t \times Private Employment_t

Why: Captures total wage bill (income mass), not just per-person earnings.

Use: Strong negative correlation with vacancy (especially residential and retail).



🚅 1.2 Wage Pressure Indicator

Formula:

$$\text{Wage Pressure}_t = \frac{\Delta \text{ Avg Hourly Earnings}}{\Delta \text{ Labor Productivity}}$$

Why: When wages grow faster than productivity, business margins compress \rightarrow may slow hiring \rightarrow higher office vacancy later.

Use: Lagged predictor for **office and industrial** vacancy.



🧈 1.3 Real Disposable Earnings

Formula:



$$ext{Real Disposable Wage} = rac{ ext{Avg Hourly Earnings}}{ ext{CPI}} imes (1 - ext{Avg Tax Rate})$$

Why: Reflects purchasing power of income — the true driver of housing and retail demand.



🏭 2. Urban Economic Balance Features

Formula:

$$ext{Elasticity} = rac{\% \ \Delta Vacancy}{\% \ \Delta AvgHourlyEarnings}$$

Why: Measures how sensitive vacancy is to wage changes — used to cluster markets by income sensitivity (high in tech-driven cities like SF, low in manufacturing cities).

2.2 Local Wage Gradient (Spatial Competitiveness)

Formula:

 $\label{eq:wage_market} \mbox{Wage Gradient} = \mbox{Avg Wage} \\ \mbox{Wage}_{Market} - \mbox{Weighted Avg Wage of Nearby Markets}$

Why: Captures inter-market labor arbitrage — if neighboring cities pay much less, high-wage markets attract talent \rightarrow lower apartment vacancy but higher office cost pressure.

2.3 "Cost of Living-Adjusted Wage"

Formula:

$$Adj Wage = \frac{Avg Hourly Earnings}{Cost of Living Index}$$

Why: Measures **net real attractiveness** of a market. High adjusted wage = migration inflow $\rightarrow \downarrow$ vacancy (especially residential).

3. Financial & Inflation Transmission Features



Formula:

$$Wage-Rate\ Ratio = rac{Avg\ Hourly\ Earnings}{10-Year\ Treasury\ Rate}$$

Why: Indicates affordability vs. borrowing cost; high ratio means earnings are growing faster than debt burden \rightarrow strong demand.



Formula:

$$Wage-Inflation Gap = Wage Growth - CPI Growth$$

Why: Real wage growth positive = affordability improvement \rightarrow residential/retail vacancy falls.



3.3 Wage-Equity Market Co-movement

Formula:

Rolling correlation (4Q or 8Q) between avg_hourly_earning_private and SP500_Price_Index.

Why: Captures macro cycle synchronization — both wages and equities rise in expansions \rightarrow demand expansion across property types.



4. Behavioral & Sentiment-Based Features



Formula:

Interaction of wage growth with Consumer Confidence Index (CCI):

Optimism Index = Wage Growth \times CCI

Why: Converts wage data into psychological purchasing power; high optimism amplifies income effect on space absorption.



4.2 Wage Volatility (Labor Uncertainty)

Rolling standard deviation of earnings growth over last 8 quarters.

Why: Wage instability \rightarrow cautious consumers and firms \rightarrow higher vacancy risk.

4.3 Structural Break Feature

Idea:

Include a **dummy or spline** when wages structurally shift (e.g., post-COVID wage surge).

Why: Detect regime changes — vacancy elasticity differs across economic regimes.

5. Industry-Level Spillover Features

🚅 5.1 Wage-to-Office Employment Ratio

 $Ratio = \frac{Avg \ Hourly \ Earnings}{Office-Using \ Employment \ Share}$

Why: High wages + high office sector share = strong office demand $\rightarrow \downarrow$ vacancy.



🚅 5.2 Wage × Industrial Employment

Composite = Avg Hourly Earnings \times Manufacturing Employment

Why: Measures industrial output capacity — useful for warehouse/industrial vacancy models.

5.3 Service Sector Wage Premium

$$Premium = \frac{Private\ Sector\ Wages}{Public\ Sector\ Wages}$$

Why: Reflects market dynamism — private sector strength correlates with commercial real estate activity.



👀 6. Income Distribution & Inequality Features



🚅 6.1 Gini-Adjusted Earnings

If income distribution data is available:

$$AdjEarnings = AvgEarnings \times (1 - Gini)$$

Why: Controls for inequality — markets with more equitable income distribution show steadier demand.

🚅 6.2 Median-to-Mean Earnings Ratio

Why: Detects skew — high skew markets (tech hubs) may have volatile vacancy due to concentrated wage earners.



🧮 7. Market Competitiveness and Attractiveness Features

Feature	Formula / Description	Real Estate Implication
Wage Momentum vs. Vacancy Momentum	Correlation of wage growth with change in vacancy	Measures demand tightness
Relative Wage Position	Market's rank in national wage percentile	High-rank markets attract population inflows
Migration-Weighted Wage Index	wages * net migration	Strong predictor of vacancy in apartments
Employment Elasticity of Wages	Regression slope of employment vs wage	Markets with responsive wage–employment dynamics have faster real estate absorption cycles

8. Machine Learning-Driven Transformations (Used in Moody's Labs & BlackRock RE Data Science)

Technique	Description	How It's Used
Dynamic Factor Model (DFM)	Combine multiple labor metrics (earnings, employment, participation) \rightarrow 1 latent "Labor Heat" feature	Input into multi-sector vacancy model
Elasticity-Encoded Features	Estimate local β from regression vacancy ~ wages per market; use β as a meta-feature	Captures market sensitivity directly
Time-Varying Coefficients (Kalman Filter)	Model changing wage-vacancy relationship over time	Detect evolving economic regimes
Residual Wage Shock	Actual wage – predicted wage from trend model	Measures "unexpected wage shock" impacting vacancy

9. Example: Advanced Pipeline (High-Level Pseudocode)

```
# Market-level engineered features
df['labor_income_index'] = df['avg_hourly_earning_private'] * df['employment_private']
df['real_wage'] = df['avg_hourly_earning_private'] / df['CPI']
df['wage_inflation_gap'] = df['wage_qoq'] - df['CPI_qoq']
df['wage_interest_ratio'] = df['avg_hourly_earning_private'] / df['ten_year_treasury_rate']
df['wage_optimism_index'] = df['wage_qoq'] * df['consumer_confidence_index']
df['market_relative_wage'] = df['avg_hourly_earning_private'] / df.groupby('quarter')
['avg_hourly_earning_private'].transform('mean')
df['wage_volatility_8q'] = df.groupby('market')
['wage_qoq'].rolling(8).std().reset_index(level=0, drop=True)
```

10. What the Top Firms Actually Do (Industry Practices)

Firm	Real Estate Analytics Use Case	Wage-Driven Feature Example
Moody's CRE	Vacancy & rent forecasting models	Real disposable income × population
CBRE Econometric Advisors	Demand forecasting	Wage-to-rent ratio & wage-to-productivity gap
BlackRock Real Estate AI	Market clustering	Dynamic factor of labor market indicators
CoStar Analytics	Submarket scoring	Wage gradient (market vs peers)
JLL Research	Urban migration prediction	Cost-of-living-adjusted wage competitiveness

Would you like me to now narrow these down to a **ranked shortlist (Top 10 most effective)** based on real-world predictive performance seen in institutional-grade vacancy models (e.g., residential vs office vs retail)?

That'll help you pick the highest-value ones to actually implement next.



So me some more features engineering with average_hourly_earning_private.

Perfect — let's **go even deeper** into *industry-grade*, creative, and data-science-driven feature engineering specifically around

average hourly earning private

for real estate modeling (especially vacancy prediction).

We'll now look at features used or inspired by top-tier economic research, institutional RE firms (Moody's, CBRE, CoStar, BlackRock), and macro-labor econometrics — going beyond even what we covered earlier.



1. Macro-Cycle Interaction Features



🚅 1.1 Wage-Business Cycle Alignment Index

Formula:

Rolling correlation (8 or 12 quarters) between wage growth and GDP growth.

Intuition: Measures whether wage growth is pro-cyclical (moving with economy).

- High correlation → expansion phase → ↓ vacancy
- Low or negative → wage rigidity or downturn → ↑ vacancy

Formula:

Wage Growth - (-∆ Unemployment Rate)

Why: When wages rise while unemployment also rises → cost-push inflation without job growth (bad for

Use: Leading indicator for **office** and **retail** vacancy.

🚅 1.2 Wage–Unemployment Divergence



💤 1.3 Wage Shock Index

Formula:

Actual Wage Growth - Predicted Wage Growth (predicted from long-term trend or ARIMA model). Why: Captures unexpected shifts in wage trends (supply shocks).

- Positive shock = more spending → ↓ vacancy
- Negative shock = stagnation → ↑ vacancy



2. Labor Income Distribution & Quality Metrics

🚅 2.1 Wage Dispersion (Inequality)

If you have quantile data (P90/P10):



$$\text{Dispersion} = \frac{P90}{P10}$$

Why: High inequality reduces overall purchasing stability → more vacancy risk in **mass-market** residential.

🚅 2.2 High-Wage Employment Share

$$Share = \frac{Employment \ in \ High-Wage \ Industries}{Total \ Employment}$$

Why: Reflects market resilience — more high-income workers = stable apartment and office demand.

🚅 2.3 Wage-Skill Premium

$$Premium = \frac{Avg \ Wage \ in \ Skilled \ Services}{Avg \ Wage \ in \ Manufacturing}$$

Why: Measures structural shift toward knowledge economy — correlated with office space absorption.

🏫 3. Housing Market Interaction Features

🚅 3.1 Wage-to-Mortgage-Cost Ratio

$$\text{Ratio} = \frac{\text{Avg Hourly Earnings}}{\text{Mortgage Rate} \times \text{Home Price Index}}$$

Why: Captures affordability power of income vs financing cost.

• High ratio \rightarrow improved buying power $\rightarrow \downarrow$ apartment vacancy.

🚅 3.2 Wage-to-Rent Differential (Pressure on Ownership Shift)

$$Diff = Wage Growth - Rent Growth$$

Why: If wages rise faster than rent \rightarrow renters can afford homes \rightarrow potential rise in vacancy for rental markets.

🚅 3.3 Housing Absorption Power Index

$$Absorption \ Index = \frac{Avg \ Hourly \ Earnings \times Population \ Growth}{Housing \ Stock \ Growth}$$

Why: Encapsulates income + demographics + supply \rightarrow directly predicts residential vacancy.

🤪 4. Sectoral Sensitivity & Linkage Features

🧈 4.1 Wage–Retail Coupling

Coupling = Correlation(Wages, RetailSales, rolling window = 4)

Why: High coupling = income translates quickly to consumption \rightarrow low **retail vacancy**.

💤 4.2 Wage × Service Employment

Feature: Avg Hourly Earning * Employment in Services

Why: Service-heavy regions turn wages into demand quickly (important for retail and office sectors).

🚅 4.3 Wage–Technology Intensity Index

Tech Intensity = Wage Growth in Tech Industries/Total Wage Growth

Why: Captures presence of innovation-driven job markets — crucial for office space forecasting.

5. Labor-Cost Pressure Features (Supply-Side)

5.1 Wage Cost Index for Developers

$$WageCostIndex = \frac{ConstructionWageGrowth}{Avg\ Hourly\ Earnings}$$

Why: High labor cost \rightarrow higher construction costs \rightarrow limited new supply $\rightarrow \downarrow$ future vacancy.

№ 5.2 Wage-Material Cost Divergence

 $Divergence = Wage\ Growth - Construction Material Cost\ Growth$

Why: If wages rise slower than material cost, construction margins shrink \rightarrow reduced new projects $\rightarrow \downarrow$ vacancy risk later.

5.3 Employment Cost Index Spread

 $Spread = Avg\; Hourly\; Earnings\; Growth - Employment CostIndex\; Growth$

Why: Measures gap between private wages and overall labor costs (signals business strain).

6. Spatial Economic Mobility Features

🚅 6.1 Wage Attraction Index

 $Attraction = (Market Wage - National Average) \times Net Migration$

Why: High-wage markets with net inflow = strong population magnet \rightarrow lower vacancy.

🚅 6.2 Wage Gradient Momentum

Compute slope of local wage difference vs distance to major city (e.g., distance from NYC). Why: Captures suburban spillover effects — rising wages spreading out \rightarrow regional vacancy tightening.

🚅 6.3 Wage–Commuter Affordability Index

$$Index = \frac{Avg \; Hourly \; Earnings \; in \; Suburb}{Transport \; Cost \; to \; City}$$

Why: When suburban wage power > commute cost \rightarrow migration to suburbs \rightarrow rising suburban demand.

🧮 7. Dynamic & Nonlinear Transformations

Туре	Description	Purpose
Polynomial terms	E^2, E^3	Capture non-linear wage-vacancy effects
Log-transform	log(E)	Interpret as % changes
Interaction with time trend	E * Quarter_Index	Capture structural time shifts (like post- pandemic regime)
Quantile encoding	Rank wage level within national distribution	Used for categorical comparisons between markets

8. Temporal Dynamics (Advanced Time-Series Features)

Feature	Explanation	
Wage Momentum (Derivative)	First derivative of wage growth rate over time	
Wage Acceleration (Second Derivative)	Detect inflection points (when wage growth starts slowing or speeding)	
Rolling Skewness	Detect asymmetric wage growth — fast rises, slow drops	
Entropy of Wage Series	Measures unpredictability — high entropy = volatile labor market \rightarrow higher vacancy risk	

🧖 9. ML-Driven or Derived Variables

ML Technique	Feature Idea	What It Captures
Dynamic Factor Analysis	Combine wages, employment, unemployment into latent "Labor Health Index"	Underlying market momentum
Clustering	Cluster markets by wage patterns \rightarrow one-hot encode cluster	Market typology feature
Residual Features	Actual Wage — Predicted Wage (trend)	Wage surprise shocks
PCA Features	Principal component of wage-related metrics	Captures joint labor strength dimension



ML Technique	Feature Idea	What It Captures
Gradient Features	Local derivative of wage w.r.t. time	Captures sudden shifts in wage trend



→ 10. Real Estate Capital Flow Interaction Features

Feature	Formula	Why Important
Wage–Cap Rate Differential	Wage Growth — CapRate	Shows income strength vs investment returns
Wage–Construction Starts Ratio	Wage Growth / ConstructionStarts	High ratio = more demand than supply
Wage–REIT Performance Co-movement	Rolling correlation with REIT Index	Reflects alignment between real estate earnings and capital market confidence

🚰 Summary: "Elite-Level" Feature Categories

Category	Feature Examples	Purpose
Macro Alignment	Wage-GDP, Wage Shock Index	Economic phase detection
Housing Demand	Wage-to-Rent, Wage-to-Mortgage	Affordability & absorption
Spatial Flow	Wage Attraction, Gradient Momentum	Migration & regional growth
Sector Linkages	Wage × Services, Wage–Retail Coupling	Sectoral demand spillover
Financial Links	Wage-Rate, Wage-CapRate	Capital cost vs income power
Nonlinear Dynamics	Entropy, Acceleration, Polynomial	Detect nonlinearities
ML-Enhanced	Latent factors, Residual shocks	Hidden wage-vacancy drivers

Would you like me to now:

- 1. Rank these features by their predictive power (as typically observed across residential, office, retail, industrial sectors)?
- 2. **Show a visual causal chain** diagram how wage dynamics cascade through economic channels into vacancy (labor \rightarrow income \rightarrow demand \rightarrow absorption \rightarrow vacancy)?