

# CCI and Vacancy Rates



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 `consumer_confidence_index` and response `vacancy`. Help me understand how it works

Perfect — this is an important and very relevant feature in real estate modeling. Let's unpack **how the Consumer Confidence Index (CCI)** works and how it connects to **vacancy rates**, especially in a real estate context.

## Step 1. What is the Consumer Confidence Index (CCI)?

The **Consumer Confidence Index** measures how optimistic or pessimistic consumers feel about the **state of the economy** and their **personal financial situation**.

- **High CCI** → Consumers feel confident about their job, income, and economic stability. They tend to **spend more** (on goods, housing, retail, etc.).
- **Low CCI** → Consumers are uncertain or pessimistic. They **cut back spending**, delay big purchases, or avoid moving homes/offices.

The CCI is published monthly by organizations like the **Conference Board (U.S.)** and captures consumer sentiment from survey data.

## Step 2. The relationship between CCI and Vacancy Rates

Vacancy rate is the percentage of available (unoccupied) real estate in a given market or sector (e.g., apartment, office, retail).

Here's how CCI can affect **vacancy** across property types:

Property Type	When CCI ↑ (confidence high)	When CCI ↓ (confidence low)
<b>Apartments / Housing</b>	More home-buying → may reduce rental demand slightly (raising rental vacancy). But overall, healthy job market keeps vacancies stable.	Renters delay moves, less mobility, potential rise in vacancies if job losses happen.
<b>Offices</b>	Businesses expand, hire more → demand for office space ↑ → vacancies ↓	Firms freeze hiring or downsize → demand for office space ↓ → vacancies ↑
<b>Retail</b>	Consumers spend more → stores expand → retail occupancy ↑ → vacancy ↓	Consumers spend less → retail closures ↑ → vacancy ↑
<b>Industrial / Logistics</b>	Higher consumer demand drives goods movement → warehouse demand ↑ → vacancy ↓	Lower spending → less demand for warehousing → vacancy ↑

 **Overall:**

- CCI has an **inverse relationship** with vacancy:  
When confidence is high, vacancy usually **drops**.  
When confidence is low, vacancy **rises**.

## Step 3. Why this matters for your model

For forecasting or causal analysis, CCI acts as a **leading indicator**:

- It reflects **future consumer spending** and **business expansion intentions**.
- Real estate activity (construction, leasing, occupancy) **lags behind** changes in confidence.

That means:

A fall in CCI today might predict **higher vacancy in the next few quarters**, not immediately.

## Step 4. Feature Engineering Ideas with `consumer_confidence_index`

To capture these dynamics, you can engineer features like:

### ◆ Lagged Features

- `CCI_lag_1`, `CCI_lag_2`, `CCI_lag_3` → 1-, 2-, or 3-quarter lag.
  - Captures delayed response of vacancy to consumer sentiment changes.

### ◆ Rolling Statistics

- `CCI_rolling_mean_4` → 4-quarter moving average (smooths out noise, captures trend).
- `CCI_rolling_std_4` → volatility of consumer sentiment (high volatility = uncertainty = potential vacancy risk).

### ◆ Growth and Momentum

- $CCI\_change\_qoq = (CCI\_t - CCI\_t-1) / CCI\_t-1$
- $CCI\_trend = CCI\_t - CCI\_rolling\_mean\_4$ 
  - Captures whether sentiment is improving or worsening relative to its long-term trend.

### ◆ Interaction Features

Combine with other variables:

- `CCI * Employment` → measures confidence supported by actual job strength.
- `CCI * Retail_Sales` → tests how confidence translates into spending behavior.
- `CCI * Inflation` → shows real confidence after adjusting for price pressure.

## ◆ Categorical Conversion

If your model benefits from thresholds:

- `High_Confidence = 1 if CCI > 100 else 0`
- Useful to separate optimistic vs pessimistic quarters.



## Step 5. Typical Expectation in Modeling

In a regression or ML context:

$$Vacancy_t = \beta_0 + \beta_1 * CCI_{t-k} + \beta_2 * Employment + \beta_3 * GDP + \dots + \varepsilon$$

Usually:

- $\beta_1 < 0 \rightarrow$  higher CCI leads to lower vacancy after some lag  $k$ .



## Step 6. Validation Ideas

You can **validate empirically**:

1. Compute the **correlation between CCI and vacancy** (lagged by 1–4 quarters).
2. Use **Granger causality tests** to check if CCI helps predict future vacancy.
3. Plot **CCI vs Vacancy (shifted by lags)** for visual intuition — often, dips in CCI precede rises in vacancy.

If you want, I can show you **advanced, real-estate-style feature engineering** used by top firms (e.g., REIS, Moody's Analytics, CBRE) that go beyond standard lags — like **sentiment shock detection**, **CCI dispersion indices**, or **business-cycle alignment**.

Would you like me to show those next?



Help with out of box features engineering for this column.

Excellent — since you're asking for **out-of-the-box (top-tier, industry-level)** feature engineering ideas for the `consumer_confidence_index` in a **real estate vacancy modeling context**, let's go beyond the standard lags and rolling means.

We'll explore **creative, high-signal, cross-domain, and macro-linked** features that top research teams (like Moody's, CBRE, JLL, and BlackRock real assets analytics) use.



## 1. Behavioral & Sentiment-Derived Features

These capture **how consumer mood transitions** rather than just levels.

## ◆ Sentiment Momentum

- $CCI\_momentum = CCI\_t - CCI_{\{t-4\}}$   
→ Measures change over the past year.  
A large drop indicates consumers are rapidly losing optimism — often a precursor to higher vacancies.

## ◆ Sentiment Shock Indicator

- $CCI\_shock = 1 \text{ if } (CCI\_t - CCI_{\{t-1\}}) < -\Delta\_threshold \text{ else } 0$   
→ Binary variable marking sudden drops beyond a threshold (e.g., -5 points QoQ).  
Real estate reacts sharply to sentiment shocks.

## ◆ Sentiment Recovery Velocity

- $CCI\_recovery\_speed = (CCI\_t - \min(CCI_{\{t-4:t\}})) / (\max(CCI_{\{t-4:t\}}) - \min(CCI_{\{t-4:t\}}))$   
→ Captures how fast confidence rebounds after a dip — useful in post-recession recovery cycles.



## 2. Business Cycle Alignment Features

CCI often mirrors the **macro cycle**, so tie it to phases of expansion/contraction.

### ◆ Economic Regime Classification

- Cluster or tag quarters into “expansion”, “early slowdown”, “recession”, “recovery” using rolling CCI trends and GDP growth.  
→ Encode  $Cycle\_Phase \in \{Expansion, Slowdown, Recession, Recovery\}$ .

### ◆ Business Cycle Synchronization

- $CCI\_phase\_shift = \text{lag when correlation}(CCI, GDP\_growth) \text{ is highest}$   
→ Quantifies whether confidence leads or lags real activity in your market.



## 3. Cross-Market and Sector Sensitivity Features

Different property types react differently to consumer confidence.

### ◆ Confidence-Elastic Vacancy Proxy

For each property type or market:

$$Elasticity = \frac{\% \Delta Vacancy}{\% \Delta CCI}$$

→ Use as a market-specific meta-feature capturing how sensitive vacancy is to sentiment shifts.

### ◆ Weighted Regional Confidence Index

If you have multiple markets:

$$CCI_{regional\_weighted} = \sum_i (CCI_i * Market\_GDP_i) / \sum_i Market\_GDP_i$$

→ Represents the macro sentiment weighted by economic influence of each region.

## 4. Income & Spending Interaction Features

Confidence alone isn't enough — link it with **income, wages, inflation, and consumption**.

### ◆ Real Confidence Index

$$CCI_{real} = \frac{CCI_t}{Inflation_t}$$

→ Adjusts sentiment for inflation pressure — high confidence amid inflation may not mean real spending.

### ◆ Spending-Backed Confidence

$$CCI_{spending\_weighted} = CCI_t * Retail\_Sales\_Growth_t$$

→ Reflects “confidence backed by actual behavior” — often has stronger predictive power for retail and apartment vacancy.

### ◆ Income-Confidence Interaction

$$CCI_{income\_interaction} = CCI_t * Avg\_Hourly\_Earnings_t$$

→ Tests if optimism coincides with purchasing power — high-value insight for housing and commercial demand.

## 5. Volatility and Uncertainty Metrics

Vacancy responds not just to mean confidence but to **sentiment uncertainty**.

### ◆ Rolling Volatility

$$CCI_{volatility} = StdDev(CCI_{t-4:t})$$

→ Measures instability in sentiment; high volatility = market hesitation = potential vacancy increases.

### ◆ Sentiment Dispersion Index

Use survey-level data (if available) or synthesize:

$$CCI_{dispersion} = |CCI_{Present} - CCI_{Expectations}|$$

→ Captures mismatch between current optimism and future expectations — predictive of turning points.



## 6. Lagged, Nonlinear, and Regime-Specific Effects

Vacancy's response to CCI is rarely linear.

### ◆ Nonlinear Transformations

- `CCI_squared` → captures diminishing or accelerating impact at extremes.
- `CCI_log_change` =  $\log(\text{CCI}_t / \text{CCI}_{\{t-1\}})$
- `CCI_rel_to_trend` =  $\text{CCI}_t / \text{RollingMean}(\text{CCI}, 8)$  — above or below long-run trend.

### ◆ Lag Structure Features

Use distributed lags:

$$\text{CCI\_weighted\_lag} = 0.5 * \text{CCI}_{t-1} + 0.3 * \text{CCI}_{t-2} + 0.2 * \text{CCI}_{t-3}$$

→ Weighted lag to represent gradual adjustment of leasing decisions to sentiment.



## 7. Composite Macro Index

Blend confidence with other leading indicators:

$$\text{Consumer\_Activity\_Composite} = z(\text{CCI}) + z(\text{Retail\_Sales}) + z(\text{Employment})$$

→ A macro-composite that reflects consumer-driven economic strength — better than CCI alone.



## 8. Shock-Response Features (Dynamic)

Inspired by econometrics and macro forecasting.

### ◆ Sentiment Shock Impact Variable

For each shock (drop > certain threshold):

- Compute average vacancy change in next 2 quarters → model as shock response parameter.

### ◆ Lagged Cross-Correlation Strength

Compute the correlation between lagged CCI and vacancy over a moving window (e.g., 12 quarters).

- `Corr_CCI_Vacancy_Lag` → time-varying relationship strength.  
Detects how predictive CCI currently is — very advanced feature.

## 9. Contextual & Structural Features

- $CCI\_relative\_to\_national\_avg = CCI\_market / CCI\_USA$   
→ Local optimism compared to national mood.
- $CCI\_gap = CCI\_current - CCI\_10yr\_avg$   
→ Measures deviation from structural normal.
- $CCI\_phase\_slope = slope(CCI\_t-4:t) \rightarrow$  short-term trend direction.

## 10. Hybrid Predictors Used by Analysts

Large research firms often use:

- $CCI\_shock\_score = (CCI\_change\_qoq / CCI\_volatility\_4q) \rightarrow$  normalized drop magnitude.
- $CCI\_consistency\_score = \% \text{ of last 4 quarters with } CCI > 100 \rightarrow$  long-term optimism persistence.
- $CCI\_reversal\_flag = 1 \text{ if } sign(CCI\_change\_qoq) \neq sign(CCI\_change\_qoq\_prev) \rightarrow$  captures turning points in sentiment cycles.

Would you like me to **tailor these feature engineering ideas by property type** (like office, retail, apartment, industrial)?

The sensitivity and useful CCI transformations differ significantly across them — e.g., **office responds to corporate confidence, retail to consumer spending**, etc.



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

Ah! Got it — you want **real estate industry-level, cutting-edge feature engineering** for **consumer\_confidence\_index** (CCI) specifically **as it relates to vacancy forecasting**. Let's go beyond the usual lags and rolling averages and focus on **"top-tier, practitioner-level" approaches** that Moody's, CBRE, JLL, and REIS use.

## 1 Macro-Adjusted Confidence Features

Industry leaders don't just use raw CCI—they adjust it to **real economic conditions**.

- **Inflation-Adjusted Confidence (Real CCI):**

$$CCI_{real} = \frac{CCI}{CPI\_growth}$$

Captures *true purchasing power sentiment* — high nominal confidence with high inflation might still reduce demand.

- **Employment-Weighted Confidence:**

$$CCI_{emp\_weighted} = CCI \times \text{Employment\_Index}$$

Only counts optimism that aligns with actual employment trends — highly predictive for office and multifamily vacancies.

- **Income-Adjusted Confidence:**

Multiply CCI by average household income or wages (`Avg_Hourly_Earnings_Private`) — reflects *spending capacity*, not just optimism.

## 2 Lagged & Distributed Effect Features

Vacancy reacts to confidence *with delay*. Top firms capture this systematically:

- **Distributed Lag CCI:**

$$CCI_{weighted\_lag} = 0.5 * CCI_{t-1} + 0.3 * CCI_{t-2} + 0.2 * CCI_{t-3}$$

Smoothly captures the delayed effect of consumer optimism on leasing behavior.

- **Dynamic Lag Selection:** Use cross-correlation to determine the **optimal lag per property type** — office, retail, industrial all respond differently.

## 3 Sentiment Shocks & Turning Points

Vacancy often spikes after **sharp drops in consumer confidence**, not gradual trends.

- **CCI Shock Indicator:**

Binary feature → 1 if CCI drops > X points QoQ.

Helps models detect recessionary triggers for vacancy.

- **Shock Magnitude Feature:**

$$CCI\_shock\_mag = \frac{CCI_t - CCI_{t-1}}{CCI\_volatility\_rolling4}$$

Normalized drop magnitude, capturing intensity relative to recent history.

- **Reversal Flag:** Detect when declining confidence starts rising again — signals recovery in leasing activity.

## 4 Cross-Indicator Interaction Features

Top analysts combine CCI with **other leading economic indicators**:

- **CCI × Retail Sales Growth** → Predicts retail vacancy shifts.
- **CCI × Housing Starts** → Apartment or residential vacancy insights.
- **CCI × Unemployment Rate** → Combines sentiment with actual labor risk.
- **CCI × SP500 or Business Confidence** → Office and industrial demand sensitivity to corporate optimism.



## 5 Volatility and Stability Features

Vacancy reacts not just to the *level* of confidence but also to **stability of sentiment**:

- **Rolling Volatility:** Standard deviation of CCI over 4 or 8 quarters — high volatility = market uncertainty = potential vacancy risk.
- **Dispersion Index (if survey data available):** Difference between current sentiment and expected sentiment.  
→ Large mismatch often precedes sudden lease reductions or slowdowns.

## 6 Cycle & Regime Features

Instead of raw CCI, top models capture **where we are in the economic cycle**:

- **Business Cycle Phase:** Map each quarter into **Expansion** / **Slowdown** / **Recession** / **Recovery** using CCI trends + GDP growth.
- **Cycle-Adjusted CCI:** Difference between current CCI and long-term moving average.

$$CCI_{gap} = CCI - RollingMean(CCI, 10yrs)$$

Highlights deviations from structural norm — often predictive of vacancy spikes.

- **CCi Trend Slope:** Short-term trend of CCI (e.g., last 4 quarters) to detect acceleration or deceleration in confidence.

## 7 Market-Specific & Weighted CCI

Top real estate analytics firms build **market-level composite indicators**:

- **Weighted Regional CCI:**

$$CCI_{regional} = \sum_i (CCI_i \times Market\_Size_i) / \sum_i Market\_Size_i$$

Accounts for markets with different economic weight — e.g., NYC vs Austin.

- **Property-Type Sensitivity Feature:**  
Compute historical correlation or elasticity of vacancy to CCI per property type — then use this to scale CCI for modeling.

## 8 Composite “Consumer Activity Index”

Some advanced models don’t use CCI in isolation:

$$Consumer\_Activity = z(CCI) + z(Retail\_Sales) + z(Housing\_Starts) + z(Employment)$$

- Standardizes and aggregates multiple leading indicators.

- Provides a **stronger, more predictive signal for vacancy forecasting** than any single feature.

## ✓ Summary of Top-Tier Features

Feature Category	Examples
Macro-adjusted	Inflation-adjusted CCI, Employment-weighted CCI, Income-adjusted CCI
Lagged/Distributed	Weighted lag CCI, Optimal lag per property type
Shock/Turning Points	CCI shock, normalized shock magnitude, reversal flag
Interactions	CCI × Retail Sales, CCI × Housing Starts, CCI × Unemployment
Volatility	Rolling std, dispersion index
Cycle/Regime	Business cycle phase, deviation from long-term trend, trend slope
Market-Specific	Weighted regional CCI, property-type elasticity
Composite	Multi-indicator consumer activity index

If you want, I can **make a concrete Python feature engineering plan** for `consumer_confidence_index` **ready to apply on quarterly real estate data**, including lags, shocks, volatility, interactions, and market-weighted composites.

Do you want me to do that next?



Some some more features engineering with `consumer_cinfidence_index`.

Perfect! Let's go **even deeper** and explore **more advanced and "out-of-the-box" feature engineering ideas** for `consumer_confidence_index` (CCI) in real estate vacancy modeling. These go beyond standard lags, shocks, and rolling averages — the kind top-tier real estate analytics teams use.

## 1 Higher-Order and Nonlinear Transformations

Vacancy does not always respond linearly to consumer confidence. Capture **nonlinearities**:

- `CCI_squared` → Captures accelerating impact at very high/low confidence levels.
- `CCI_cubed` → Sensitive to extreme optimism or pessimism.
- `CCI_log_change = log(CCI_t / CCI_{t-1})` → Models relative growth rather than absolute change.
- `CCI_rel_to_trend = CCI_t / RollingMean(CCI, 8)` → Measures above/below trend deviations.

## 2 Rate-of-Change & Momentum Features

How fast sentiment changes can signal market reactions:

- `CCI_qoq_change = CCI_t - CCI_{t-1}` → Quarterly change.

- $CCI\_momentum\_4q = CCI\_t - CCI_{\{t-4\}}$  → Annual change, captures cyclical momentum.
- $CCI\_acceleration = (CCI\_t - CCI_{\{t-1\}}) - (CCI_{\{t-1\}} - CCI_{\{t-2\}})$  → Detects acceleration/deceleration in consumer optimism.

### 3 Lagged Cumulative Features

Vacancy often responds to **cumulative confidence changes** over multiple quarters:

- $CCI\_cum4q = \text{sum}(CCI_{\{t-1\}}, \dots, CCI_{\{t-4\}})$  → Total sentiment over last year.
- $CCI\_weighted\_lag = 0.5*CCI_{\{t-1\}} + 0.3*CCI_{\{t-2\}} + 0.2*CCI_{\{t-3\}}$  → Weighted influence of past quarters.

### 4 Volatility & Uncertainty Features

Top analysts model the **stability of confidence**, not just the level:

- $CCI\_rolling\_std4 = \text{StdDev}(CCI_{\{t-3:t\}})$  → Volatility over last 4 quarters.
- $CCI\_rolling\_cv4 = \text{StdDev} / \text{Mean}$  → Coefficient of variation; high = uncertainty.
- $CCI\_dispersion = \text{abs}(CCI\_current - CCI\_expected)$  → Measures mismatch between current and expected sentiment (if survey breakdowns available).

### 5 Shock Detection Features

Vacancy reacts to **sharp declines or reversals** in confidence:

- $CCI\_shock\_flag = 1 \text{ if } CCI\_drop\_qoq > \text{threshold} \text{ else } 0$
- $CCI\_shock\_magnitude = CCI\_drop\_qoq / CCI\_rolling\_std4$  → Normalized intensity of drop.
- $CCI\_reversal\_flag = 1 \text{ if } \text{sign}(CCI\_change\_qoq) \neq \text{sign}(CCI\_change\_qoq\_prev) \text{ else } 0$  → Captures turning points.

### 6 Interaction Features with Macro & Market Variables

Confidence alone isn't enough. Combine with **other leading indicators**:

- $CCI * \text{Retail\_Sales\_Growth}$  → Predict retail vacancy.
- $CCI * \text{Avg\_Hourly\_Earnings\_Private}$  → Captures spending power.
- $CCI * \text{Unemployment\_Rate}$  → Models combined sentiment & risk exposure.
- $CCI * \text{Housing\_Starts}$  → Apartment vacancy sensitivity.

### 7 Cycle & Trend Features

Place confidence in the context of economic cycles:

- `CCI_phase = categorical(Expansion, Slowdown, Recession, Recovery)` → Map quarters using rolling CCI + GDP growth.
- `CCI_trend_slope = slope(CCI_{t-3:t})` → Short-term acceleration/deceleration.
- `CCI_gap_longterm = CCI_t - RollingMean(CCI, 40q)` → Deviation from long-term norm.

## 8 Market-Specific Features

Real estate response varies by market:

- `CCI_weighted_by_market_size = sum(CCI_i * Market_GDP_i) / sum(Market_GDP_i)` → Larger markets have more influence.
- `CCI_property_elasticity = historical %ΔVacancy / %ΔCCI` → Scales CCI by property-type sensitivity.
- `CCI_local_relative = CCI_market / CCI_national` → Local optimism compared to national average.

## 9 Composite Features

Combine multiple indicators to strengthen signal:

- `Consumer_Activity_Index = z(CCI) + z(Retail_Sales) + z(Housing_Starts) + z(Employment)`
- `CCI_macro_factor = PCA([CCI, Retail_Sales, GDP_growth, Employment])` → Dimensionality reduction for multi-indicator signal.

## 10 Advanced Time-Series Features

- `CCI_autocorr_lag1 = autocorr(CCI, lag=1)` → How persistent is confidence?
- `CCI_cross_corr_vacancy = correlation(CCI_lagged, vacancy)` → Measures predictive power over time.
- `CCI_seasonal_adjusted = CCI - seasonal_mean_quarter` → Removes seasonal bias.

If you want, I can **create a concrete Python feature engineering script** for `consumer_confidence_index` that generates **all these top-notch features automatically** for quarterly real estate vacancy modeling. This would include lags, shocks, momentum, volatility, interactions, composites, and cycle-adjusted features.

Do you want me to do that?