Characteristics Space

Foundations • The Valuation Engineer

Bert Craytor• May 29, 2026
Version 0.1.0-draft

Formal definition. The characteristics space of a class of heterogeneous goods is the vector space whose coordinates index the attributes that buyers value, and in which each good is represented by the vector of its attribute levels.

Intuitive framing. The previous entry established that a heterogeneous good is a distinct bundle of characteristics. The natural next question is: how should we describe and compare such bundles formally? The answer, due to Lancaster (1966), is to treat each good as a point in a vector space whose axes are the characteristics buyers care about.

In this representation, a Pacifica residential property is no longer "a house in Park Pacifica" but a vector with coordinates for gross living area, lot size, view, condition, school district, distance to the ocean, and so on. Two properties that share all of these attribute values would occupy the same point in characteristics space; properties that differ in any attribute occupy different points. The geometry of the space encodes how close or how different two properties are, in a way that captures all the dimensions on which buyers make distinctions.

The dimensionality of the space depends on what attributes are relevant to the market under study. For a typical residential appraisal, the relevant space might have a dozen dimensions: physical (GLA, lot, beds, baths, age, quality, condition), locational (school district, view, proximity to amenities or disamenities), and legal/economic (zoning, easements). For a commercial appraisal, the relevant dimensions shift toward cash-flow attributes (lease structure, tenant quality, occupancy), but the underlying geometric idea is the same.

What makes the characteristics space useful is that it converts a qualitative practitioner's intuition ("these comps are similar to the subject") into a quantitative claim ("these comps are close to the subject in characteristics space"). Once distance is defined, every sales comparison decision can be examined as a statement about location in the space.

Where appraisers encounter it. Appraisers work in characteristics space constantly, usually without naming it. Every line of a sales comparison grid is a coordinate axis; every comp is a row vector of coordinate values; every adjustment is an attempt to move along one axis while holding the others fixed.

The eight Pacifica comparables we will use throughout this issue illustrate the construction directly. Each comp is a point in a four-dimensional characteristics space spanned by GLA, lot size, view, and condition:

Comp	GLA (sf)	Lot (sf)	View	Cond.	Sale price
A	1,650	9,000	yes	good	$1,425,000
B	1,650	7,200	no	good	$1,280,000
C	1,850	7,500	no	good	$1,407,500
D	1,450	7,000	no	good	$1,155,000
E	1,700	8,000	yes	excellent	$1,470,000
F	1,550	7,100	no	average	$1,177,500
G	1,700	10,500	no	good	$1,392,500
H	1,600	7,300	no	good	$1,312,500

Comps A and B from the previous entry are the first two rows; the remaining six expand the example in ways we will exploit through the rest of the issue.

Each comp is a point in $\mathbb{R}^4$:

- Comp A is at $(1650, 9000, 1, 2)$.
- Comp B is at $(1650, 7200, 0, 2)$.
- Comp G is at $(1700, 10500, 0, 2)$.

The view and condition coordinates are encoded numerically (view as $\{0, 1\}$, condition as $\{1, 2, 3\}$); the price is *not* a coordinate of the characteristics space but a function of position within it.

Why it matters for defensibility. The characteristics space framework sharpens what comp selection actually is. A comparable is not just "a similar property" but a point in the relevant region of characteristics space close to the subject. The closer a comp is to the subject in the space, the smaller the adjustments required and the more reliable the indicated value becomes.

This reframing exposes several practice questions that the informal grid obscures:

- *Which dimensions matter?* The choice of axes is a modeling decision. Including "view" as an axis only makes sense if the market prices view distinctly; including "Feng Shui orientation" only makes sense if a market segment prices it. Defensible appraisal requires defending the axes, not just the values along them.
- *What distance metric applies?* Two comps may differ by $200,000 in price because they are far apart in characteristics space, but the distance computation requires a scaling for each axis. Untransformed Euclidean distance treats one square foot of GLA as equivalent to one square foot of lot, which the market does not.
- *Are the comps in a region where the price function is well-behaved?* If the subject sits in a sparsely-sampled corner of characteristics space, no comp lies close to it and the value indication is necessarily extrapolative.

Worked appraisal example. The eight Pacifica comps occupy a relatively compact region of the four-dimensional characteristics space. Consider the question of which comp is "closest" to a hypothetical subject property at $(1{,}700, 7{,}500, 1, 2)$ — 1,700 sf, 7,500 sf lot, view, good condition.

If we naively use Euclidean distance with unscaled axes, the lot dimension dominates because its values (7,000–10,500) are numerically much larger than view (0–1) or condition (1–3). Distances to the subject:

Comp	Distance (unscaled Euclidean)	Notes
A	1,501	view matches; lot dominates
B	308	view differs; lot close
C	230	view differs; closest by raw distance
D	559	view differs; GLA off
E	502	view matches; condition differs
F	422	view and condition differ
G	3,000	view differs; lot far away
H	224	view differs; closest of all

By raw distance, Comp H is the nearest neighbor. But Comp H has no view and the subject does. The naive metric has misled us: the lot dimension's numerical scale crowded out the categorical view dimension, which actually matters more to the price.

A defensible distance metric would scale each axis by its implicit price — roughly, by how many dollars one unit of variation in that axis is worth. Once the axes are scaled by implicit price, view (worth ˜$100,000) is no longer dominated by lot (˜$25 per square foot). Comp A, which matches the subject on view, then emerges as the appropriate primary comparable.

This is precisely what the implicit-price machinery developed in the following entries provides: a principled scaling of the characteristics space that turns "closeness" from a geometric guess into a market-derived computation.

Cross-references: heterogeneous good; *hedonic price function*; *implicit price.*