Differences between 2015 and 2020 datasets

* Variable name:
  + CDD 30 year
  + HDD 30 year
  + IECC climate zone
  + In 2020 number of adults divided between two variables, 18-64 and 65+
* Variable type
  + Division encoded as numeric in 2015 but character in 2020
* Variable values
  + Different income brackets in 2015 and 2020. Group brackets to have the following breakdown:

|  |  |
| --- | --- |
| 1 | < $20,000 |
| 2 | $20,000 - $39,999 |
| 3 | $40,000 - $59,999 |
| 4 | $60,000 - $99,999 |
| 5 | > $100,000 |

* + 2020 has separates unemployed and retired person. Revert to 2015 breakdown and values
  + WINDOWS had different coding
  + Recode cooling systems in 2015 and 2020 to have following breakdown:

|  |  |
| --- | --- |
| 1 | Central system only |
| 2 | Individual/portable system only |
| 3 | Multiple system types (central, individual portable systems, or evaporative/swamp cooler) |
| 6 | Evaporative or swamp cooler only |
| -2 | NA |

* + Recode other heating type in 2015 to have value of 99 instead of 21
  + Add level 21 to heating equipment type to mean multiple system types
  + Recode values in AGECENAC and WWACAGE to match breakdown in 2020 dataset. Take the maximum of these two variables and combine into one variable, ACEQUIPAGE