In [1]:

**import** **pandas**

In [2]:

**import** **requests**

In [3]:

url = 'https://finalfantasy.fandom.com/wiki/Final\_Fantasy\_VII\_armor'

In [4]:

r = requests.get(url)

In [5]:

dfs = pandas.read\_html(r.text, attrs={'class': 'full-width FFVII article-table'})

In [6]:

df = dfs[0]

In [7]:

df

Out[7]:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Armor** | **Def** | **Def%** | **MDef** | **MDef%** | **Materia** | **Cost** | **Acquire** | **Unnamed: 8** |
| **0** | Bronze Bangle | 8 | 0 | 0 | 0 | None | 100 | Initial: Cloud, Barret, Tifa and Aeris | NaN |
| **1** | Bronze Bangle | None | None | None | None | None | None | None | None |
| **2** | Iron Bangle | 10 | 0 | 2 | 0 | x1 Normal | 160 | Buy: Sector 7 Slums | NaN |
| **3** | Iron Bangle | None | None | None | None | None | None | None | None |
| **4** | Titan Bangle | 14 | 2 | 4 | 0 | x2 Normal | 280 | Buy: Sector 5 Slums, Wall MarketDrop: Air Buster | NaN |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **67** | Precious Watch | 0 | 0 | 0 | 0 | x8 Normal | 2 | Win: Chocobo racing | NaN |
| **68** | Precious Watch | None | None | None | None | None | None | None | None |
| **69** | Precious Watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch |
| **70** | Chocobracelet | 35 | 10 | 38 | 10 | x4 Normal | 2 | Win: Chocobo racing | NaN |
| **71** | Chocobracelet | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 |

72 rows × 9 columns

In [8]:

df = df.drop(['Materia', 'Acquire', 'Unnamed: 8'], axis='columns')

In [9]:

df

Out[9]:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Armor** | **Def** | **Def%** | **MDef** | **MDef%** | **Cost** |
| **0** | Bronze Bangle | 8 | 0 | 0 | 0 | 100 |
| **1** | Bronze Bangle | None | None | None | None | None |
| **2** | Iron Bangle | 10 | 0 | 2 | 0 | 160 |
| **3** | Iron Bangle | None | None | None | None | None |
| **4** | Titan Bangle | 14 | 2 | 4 | 0 | 280 |
| **...** | ... | ... | ... | ... | ... | ... |
| **67** | Precious Watch | 0 | 0 | 0 | 0 | 2 |
| **68** | Precious Watch | None | None | None | None | None |
| **69** | Precious Watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch | A gorgeous watch |
| **70** | Chocobracelet | 35 | 10 | 38 | 10 | 2 |
| **71** | Chocobracelet | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 | Dexterity +30; Luck +20 |

72 rows × 6 columns

In [10]:

df.drop\_duplicates('Armor', inplace=**True**)

In [11]:

df

Out[11]:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Armor** | **Def** | **Def%** | **MDef** | **MDef%** | **Cost** |
| **0** | Bronze Bangle | 8 | 0 | 0 | 0 | 100 |
| **2** | Iron Bangle | 10 | 0 | 2 | 0 | 160 |
| **4** | Titan Bangle | 14 | 2 | 4 | 0 | 280 |
| **6** | Mythril Armlet | 18 | 3 | 8 | 0 | 350 |
| **8** | Carbon Bangle | 27 | 3 | 14 | 0 | 800 |
| **10** | Silver Armlet | 34 | 4 | 22 | 0 | 1300 |
| **12** | Gold Armlet | 46 | 4 | 28 | 0 | 2000 |
| **14** | Diamond Bangle | 57 | 6 | 37 | 0 | 3200 |
| **16** | Crystal Bangle | 70 | 8 | 45 | 1 | 4800 |
| **18** | Platinum Bangle | 20 | 0 | 12 | 0 | 1800 |
| **20** | Rune Armlet | 43 | 5 | 24 | 0 | 3700 |
| **22** | Edincoat | 50 | 0 | 33 | 0 | 8000 |
| **24** | Wizard Bracelet | 6 | 3 | 85 | 3 | 12000 |
| **26** | Adaman Bangle | 93 | 0 | 23 | 0 | 2 |
| **28** | Gigas Armlet | 59 | 0 | 0 | 0 | 2 |
| **30** | Imperial Guard | 82 | 0 | 74 | 0 | 2 |
| **32** | Aegis Armlet[Aegis Bracelet[note 1]] | 55 | 15 | 86 | 50 | 2 |
| **34** | Fourth Bracelet | 74 | 3 | 100 | 3 | 2 |
| **36** | Warrior Bangle[Warriors Bangle[note 1]] | 96 | 0 | 21 | 0 | 1000 |
| **38** | Shinra Beta | 30 | 0 | 0 | 0 | 950 |
| **40** | Shinra Alpha | 77 | 0 | 34 | 0 | 2 |
| **42** | Four Slots(Four Slot[note 2]) | 12 | 0 | 10 | 0 | 1300 |
| **44** | Fire Armlet | 72 | 8 | 52 | 3 | 2 |
| **47** | Aurora Armlet | 76 | 8 | 54 | 3 | 4800 |
| **50** | Bolt Armlet | 74 | 8 | 55 | 3 | 2 |
| **53** | Dragon Armlet | 58 | 3 | 47 | 2 | 2 |
| **56** | Minerva Band | 60 | 8 | 57 | 0 | 2 |
| **59** | Escort Guard | 62 | 5 | 55 | 0 | 2 |
| **62** | Mystile(Mythril[note 4]) | 65 | 50 | 72 | 60 | 2 |
| **64** | Ziedrich | 100 | 15 | 98 | 18 | 2 |
| **67** | Precious Watch | 0 | 0 | 0 | 0 | 2 |
| **70** | Chocobracelet | 35 | 10 | 38 | 10 | 2 |

In [12]:

**import** **re**

In [13]:

re.sub(r'^([a-zA-Z\s]+)(\(|\[).\*$', r'\g<1>', 'Precious Watch(Mythril[note 4])')

Out[13]:

'Precious Watch'

In [14]:

df['Armor'].apply(**lambda** armor: re.sub(r'^([a-zA-Z\s]+)(\(|\[).\*$', r'\g<1>', armor))

Out[14]:

0 Bronze Bangle  
2 Iron Bangle  
4 Titan Bangle  
6 Mythril Armlet  
8 Carbon Bangle  
10 Silver Armlet  
12 Gold Armlet  
14 Diamond Bangle  
16 Crystal Bangle  
18 Platinum Bangle  
20 Rune Armlet  
22 Edincoat  
24 Wizard Bracelet  
26 Adaman Bangle  
28 Gigas Armlet  
30 Imperial Guard  
32 Aegis Armlet  
34 Fourth Bracelet  
36 Warrior Bangle  
38 Shinra Beta  
40 Shinra Alpha  
42 Four Slots  
44 Fire Armlet  
47 Aurora Armlet  
50 Bolt Armlet  
53 Dragon Armlet  
56 Minerva Band  
59 Escort Guard  
62 Mystile  
64 Ziedrich  
67 Precious Watch  
70 Chocobracelet  
Name: Armor, dtype: object

In [15]:

df['Armor'] = df['Armor'].apply(**lambda** armor: re.sub(r'^([a-zA-Z\s]+)(\(|\[).\*$', r'\g<1>', armor))

In [16]:

df

Out[16]:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Armor** | **Def** | **Def%** | **MDef** | **MDef%** | **Cost** |
| **0** | Bronze Bangle | 8 | 0 | 0 | 0 | 100 |
| **2** | Iron Bangle | 10 | 0 | 2 | 0 | 160 |
| **4** | Titan Bangle | 14 | 2 | 4 | 0 | 280 |
| **6** | Mythril Armlet | 18 | 3 | 8 | 0 | 350 |
| **8** | Carbon Bangle | 27 | 3 | 14 | 0 | 800 |
| **10** | Silver Armlet | 34 | 4 | 22 | 0 | 1300 |
| **12** | Gold Armlet | 46 | 4 | 28 | 0 | 2000 |
| **14** | Diamond Bangle | 57 | 6 | 37 | 0 | 3200 |
| **16** | Crystal Bangle | 70 | 8 | 45 | 1 | 4800 |
| **18** | Platinum Bangle | 20 | 0 | 12 | 0 | 1800 |
| **20** | Rune Armlet | 43 | 5 | 24 | 0 | 3700 |
| **22** | Edincoat | 50 | 0 | 33 | 0 | 8000 |
| **24** | Wizard Bracelet | 6 | 3 | 85 | 3 | 12000 |
| **26** | Adaman Bangle | 93 | 0 | 23 | 0 | 2 |
| **28** | Gigas Armlet | 59 | 0 | 0 | 0 | 2 |
| **30** | Imperial Guard | 82 | 0 | 74 | 0 | 2 |
| **32** | Aegis Armlet | 55 | 15 | 86 | 50 | 2 |
| **34** | Fourth Bracelet | 74 | 3 | 100 | 3 | 2 |
| **36** | Warrior Bangle | 96 | 0 | 21 | 0 | 1000 |
| **38** | Shinra Beta | 30 | 0 | 0 | 0 | 950 |
| **40** | Shinra Alpha | 77 | 0 | 34 | 0 | 2 |
| **42** | Four Slots | 12 | 0 | 10 | 0 | 1300 |
| **44** | Fire Armlet | 72 | 8 | 52 | 3 | 2 |
| **47** | Aurora Armlet | 76 | 8 | 54 | 3 | 4800 |
| **50** | Bolt Armlet | 74 | 8 | 55 | 3 | 2 |
| **53** | Dragon Armlet | 58 | 3 | 47 | 2 | 2 |
| **56** | Minerva Band | 60 | 8 | 57 | 0 | 2 |
| **59** | Escort Guard | 62 | 5 | 55 | 0 | 2 |
| **62** | Mystile | 65 | 50 | 72 | 60 | 2 |
| **64** | Ziedrich | 100 | 15 | 98 | 18 | 2 |
| **67** | Precious Watch | 0 | 0 | 0 | 0 | 2 |
| **70** | Chocobracelet | 35 | 10 | 38 | 10 | 2 |

In [17]:

df.reset\_index(inplace=**True**)

In [18]:

df

Out[18]:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **index** | **Armor** | **Def** | **Def%** | **MDef** | **MDef%** | **Cost** |
| **0** | 0 | Bronze Bangle | 8 | 0 | 0 | 0 | 100 |
| **1** | 2 | Iron Bangle | 10 | 0 | 2 | 0 | 160 |
| **2** | 4 | Titan Bangle | 14 | 2 | 4 | 0 | 280 |
| **3** | 6 | Mythril Armlet | 18 | 3 | 8 | 0 | 350 |
| **4** | 8 | Carbon Bangle | 27 | 3 | 14 | 0 | 800 |
| **5** | 10 | Silver Armlet | 34 | 4 | 22 | 0 | 1300 |
| **6** | 12 | Gold Armlet | 46 | 4 | 28 | 0 | 2000 |
| **7** | 14 | Diamond Bangle | 57 | 6 | 37 | 0 | 3200 |
| **8** | 16 | Crystal Bangle | 70 | 8 | 45 | 1 | 4800 |
| **9** | 18 | Platinum Bangle | 20 | 0 | 12 | 0 | 1800 |
| **10** | 20 | Rune Armlet | 43 | 5 | 24 | 0 | 3700 |
| **11** | 22 | Edincoat | 50 | 0 | 33 | 0 | 8000 |
| **12** | 24 | Wizard Bracelet | 6 | 3 | 85 | 3 | 12000 |
| **13** | 26 | Adaman Bangle | 93 | 0 | 23 | 0 | 2 |
| **14** | 28 | Gigas Armlet | 59 | 0 | 0 | 0 | 2 |
| **15** | 30 | Imperial Guard | 82 | 0 | 74 | 0 | 2 |
| **16** | 32 | Aegis Armlet | 55 | 15 | 86 | 50 | 2 |
| **17** | 34 | Fourth Bracelet | 74 | 3 | 100 | 3 | 2 |
| **18** | 36 | Warrior Bangle | 96 | 0 | 21 | 0 | 1000 |
| **19** | 38 | Shinra Beta | 30 | 0 | 0 | 0 | 950 |
| **20** | 40 | Shinra Alpha | 77 | 0 | 34 | 0 | 2 |
| **21** | 42 | Four Slots | 12 | 0 | 10 | 0 | 1300 |
| **22** | 44 | Fire Armlet | 72 | 8 | 52 | 3 | 2 |
| **23** | 47 | Aurora Armlet | 76 | 8 | 54 | 3 | 4800 |
| **24** | 50 | Bolt Armlet | 74 | 8 | 55 | 3 | 2 |
| **25** | 53 | Dragon Armlet | 58 | 3 | 47 | 2 | 2 |
| **26** | 56 | Minerva Band | 60 | 8 | 57 | 0 | 2 |
| **27** | 59 | Escort Guard | 62 | 5 | 55 | 0 | 2 |
| **28** | 62 | Mystile | 65 | 50 | 72 | 60 | 2 |
| **29** | 64 | Ziedrich | 100 | 15 | 98 | 18 | 2 |
| **30** | 67 | Precious Watch | 0 | 0 | 0 | 0 | 2 |
| **31** | 70 | Chocobracelet | 35 | 10 | 38 | 10 | 2 |

In [19]:

df.drop('index', axis='columns', inplace=**True**)

In [20]:

df.columns = map(str.lower, df.columns)

In [21]:

df

Out[21]:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **armor** | **def** | **def%** | **mdef** | **mdef%** | **cost** |
| **0** | Bronze Bangle | 8 | 0 | 0 | 0 | 100 |
| **1** | Iron Bangle | 10 | 0 | 2 | 0 | 160 |
| **2** | Titan Bangle | 14 | 2 | 4 | 0 | 280 |
| **3** | Mythril Armlet | 18 | 3 | 8 | 0 | 350 |
| **4** | Carbon Bangle | 27 | 3 | 14 | 0 | 800 |
| **5** | Silver Armlet | 34 | 4 | 22 | 0 | 1300 |
| **6** | Gold Armlet | 46 | 4 | 28 | 0 | 2000 |
| **7** | Diamond Bangle | 57 | 6 | 37 | 0 | 3200 |
| **8** | Crystal Bangle | 70 | 8 | 45 | 1 | 4800 |
| **9** | Platinum Bangle | 20 | 0 | 12 | 0 | 1800 |
| **10** | Rune Armlet | 43 | 5 | 24 | 0 | 3700 |
| **11** | Edincoat | 50 | 0 | 33 | 0 | 8000 |
| **12** | Wizard Bracelet | 6 | 3 | 85 | 3 | 12000 |
| **13** | Adaman Bangle | 93 | 0 | 23 | 0 | 2 |
| **14** | Gigas Armlet | 59 | 0 | 0 | 0 | 2 |
| **15** | Imperial Guard | 82 | 0 | 74 | 0 | 2 |
| **16** | Aegis Armlet | 55 | 15 | 86 | 50 | 2 |
| **17** | Fourth Bracelet | 74 | 3 | 100 | 3 | 2 |
| **18** | Warrior Bangle | 96 | 0 | 21 | 0 | 1000 |
| **19** | Shinra Beta | 30 | 0 | 0 | 0 | 950 |
| **20** | Shinra Alpha | 77 | 0 | 34 | 0 | 2 |
| **21** | Four Slots | 12 | 0 | 10 | 0 | 1300 |
| **22** | Fire Armlet | 72 | 8 | 52 | 3 | 2 |
| **23** | Aurora Armlet | 76 | 8 | 54 | 3 | 4800 |
| **24** | Bolt Armlet | 74 | 8 | 55 | 3 | 2 |
| **25** | Dragon Armlet | 58 | 3 | 47 | 2 | 2 |
| **26** | Minerva Band | 60 | 8 | 57 | 0 | 2 |
| **27** | Escort Guard | 62 | 5 | 55 | 0 | 2 |
| **28** | Mystile | 65 | 50 | 72 | 60 | 2 |
| **29** | Ziedrich | 100 | 15 | 98 | 18 | 2 |
| **30** | Precious Watch | 0 | 0 | 0 | 0 | 2 |
| **31** | Chocobracelet | 35 | 10 | 38 | 10 | 2 |

In [22]:

df.to\_csv("~/college/mobile-dev/semester\_project/data/ff7\_armors.csv")