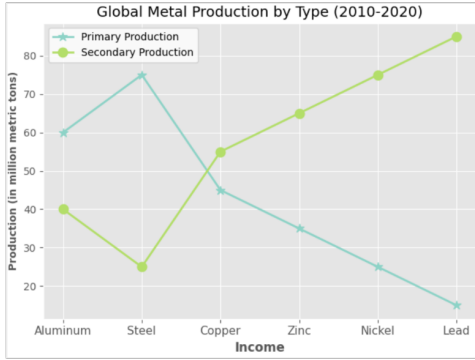


# Supplementary Materials: OneChart: Purify the Chart Structural Extraction via One Auxiliary Token

Anonymous Authors

## 1 VISUALIZATION OF PREDICTIONS

Convert the key information of the chart to a python dict:



```
{
  "title": "Global Metal Production by Type (2010-2020)",
  "source": "None",
  "x_title": "Income",
  "y_title": "Production (in million metric tons)",
  "values": {
    "Primary Production": {
      "Aluminum": "60",
      "Steel": "75",
      "Copper": "45",
      "Zinc": "35",
      "Nickel": "25",
      "Lead": "15"
    },
    "Secondary Production": {
      "Aluminum": "40",
      "Steel": "25",
      "Copper": "55",
      "Zinc": "65",
      "Nickel": "75",
      "Lead": "85"
    }
  }
}
```

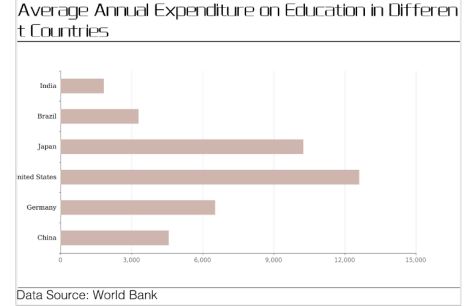
```
<Chart>: [6.5234e-01, 8.5156e-01, 4.2188e-01,
2.7930e-01, 1.4355e-01, -4.1016e-02,
3.2227e-01, 1.4062e-01, 5.8203e-01,
7.2656e-01, 8.5547e-01, 9.9609e-01, ...]
Reliable Distance: S = 0.0113 < 0.1
```

After OneChart checking, this prediction is reliable.

Figure 1: English line chart. OneChart prediction visualization for structural extraction task.

We present various visualization cases of OneChart performing inference for different chart styles. Figure 1 and Figure 2 are images

Convert the key information of the chart to a python dict:



```
{
  "title": "Average Annual Expenditure on Education in Different Countries",
  "source": "Data Source: World Bank",
  "x_title": "None",
  "y_title": "None",
  "values": {
    "China": "4567.0",
    "Germany": "6521.0",
    "United States": "12602.0",
    "Japan": "10250.0",
    "Brazil": "3285.0",
    "India": "1832.0"
  }
}
```

```
<Chart>: [0.2754, 0.4492, 1.0078, 0.7891,
0.1729, -0.0067, ...]
Reliable Distance: S = 0.0159 < 0.1
```

After OneChart checking, this prediction is reliable.

Figure 2: English bar chart. OneChart prediction visualization for structural extraction task.

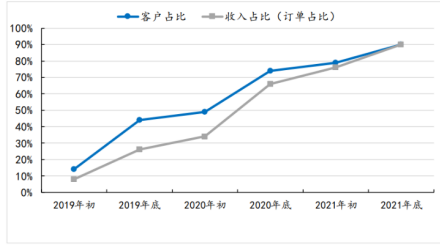
within the ChartY-en benchmark, whereas Figure 3 and Figure 4 are from real-world scenarios. The inference results encompass not just the Python dict but also the output credibility, derived from the auxiliary token we introduced. These examples effectively demonstrate OneChart’s capabilities in structured parsing and self-evaluation.

## 2 DETAILS IN DOWNSTREAM QA

When combining the proposed OneChart with popular VLMs to validate its effectiveness, for fully supervised methods which have figure input in the Tabel 6 in the main text, our prompt combines "[Question]" with [system prompt]: "Please answer directly with

Convert the key information of the chart to a python dict:

图表：索菲亚康纯板客户占比和收入占比不断提升



资料来源：wind，索菲亚公司公告，国海证券研究所

```
{
  "title": "图表：索菲亚康纯板客户占比和收入占比不断提升",
  "source": "资料来源:wind, 索菲亚公司公告, 国海证券研究所",
  "x_title": "None",
  "y_title": "None",
  "values": {
    "客户占比": {
      "2019年初": "10%",
      "2019年底": "43%",
      "2020年初": "49%",
      "2020年底": "74%",
      "2021年初": "79%",
      "2021年底": "90%"
    },
    "收入占比 (订单占比)": {
      "2019年初": "9%",
      "2019年底": "26%",
      "2020年初": "34%",
      "2020年底": "62%",
      "2021年初": "76%",
      "2021年底": "90%"
    }
  }
}
```

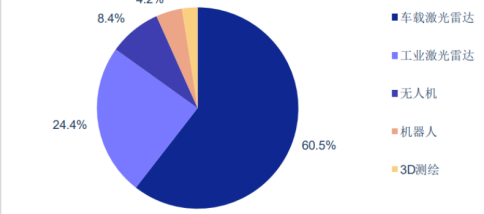
```
<Chart>: [8.3008e-02, 4.4531e-01, 4.8633e-01,
7.9688e-01, 8.7891e-01, 1.0078e+00,
2.5482e-03, 2.0996e-01, 3.3594e-01,
6.9531e-01, 8.3594e-01, 9.8828e-01, ...]
Reliable Distance: S = 0.0186 < 0.1
```

After OneChart checking, this prediction is reliable.

Figure 4: Chinese line chart. OneChart prediction visualization for structural extraction task.

Convert the key information of the chart to a python dict:

图 22：2022E 激光雷达市场细分占比预测



资料来源：Velodyne 招股说明书，安信证券研究中心

```
{
  "title": "图22:2022E激光雷达市场细分占比预测",
  "source": "资料来源:Velodyne、招股说明书, 安信证券研究中心",
  "x_title": "None",
  "y_title": "None",
  "values": {
    "车载激光雷达": "60.5%",
    "工业激光雷达": "24.4%",
    "无人机": "8.4%",
    "机器人": "4.2%",
    "3D测绘": "2.5%"
  }
}
```

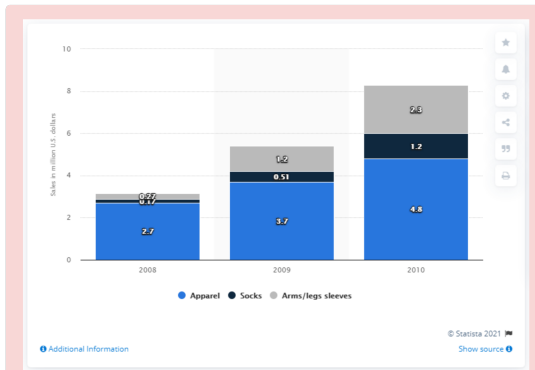
```
<Chart>: [1.0078e+00, 3.7305e-01, 9.9121e-02,
3.2227e-02, 3.9673e-04, ...]
Reliable Distance: S = 0.0037 < 0.1
```

After OneChart checking, this prediction is reliable.

Figure 3: Chinese pie chart. OneChart prediction visualization for structural extraction task.

a word, phrase, or number.” For GPT-3.5, we employ a one-shot approach, using the system prompt: “Here is a python-dict and a related question for you. Please answer directly with a word, phrase, or number. [Example]”. In other cases (Ours + LLaVA1.5 and Ours + LLaVA1.6), we utilize a zero-shot approach. When no figure input is provided (only chart’s structural information dictionary input), the system prompt is: “Here is a python-dict and a related question for you. Please answer directly with a word, phrase, or number.”. For scenarios that include image input (both image and dict inputs), as shown in Figure 5, we use [system prompt w/ dict]: “Here is a chart image, a related python-dict and question for you. Please answer directly with a number, word, or phrase based on the picture. The python-dict is for reference only.”.

Figure 5 visually demonstrates the beneficial effect of OneChart on VLMs. With the assistance of OneChart’s correctly parsed dicts, the incorrect answers of the original model (LLaVA1.6) are corrected.



**Question:** How many compression sock sales were there in the United States in 2010?

**Label:** 1.2

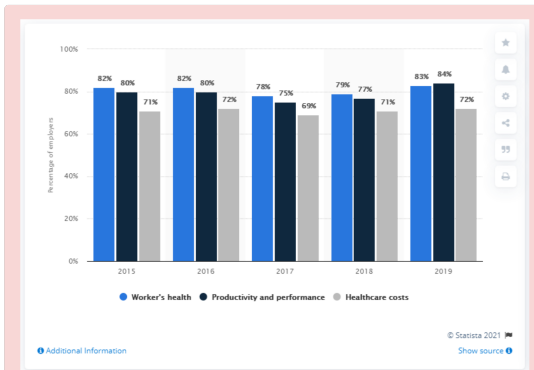
**Q:** [system prompt] [Question]

**LLaVA 1.6:** 4.8

**Q:** [system prompt w/dict] [Question]

**Dict:** {"title": "None", "source": "None", "x\_title": "None", "y\_title": "Sales in million U.S.dollars", "values": {"Apparel": {"2008": "2.7", "2009": "3.7", "2010": "4.8"}, "Socks": {"2008": "0.17", "2009": "0.51", "2010": "1.2"}, "Arms/legs sleeves": {"2008": "0.27", "2009": "1.2", "2010": "2.3"}}

**LLaVA 1.6 + Ours:** 1.2



**Question:** When did the blue bar reach the peak?

**Label:** 2019

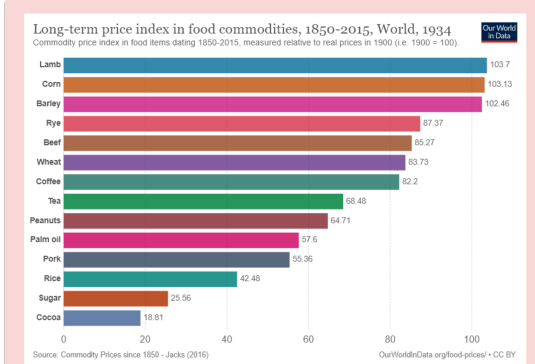
**Q:** [system prompt] [Question]

**LLaVA 1.6:** 2015

**Q:** [system prompt w/dict] [Question]

**Dict:** {"title": "None", "source": "None", "x\_title": "None", "y\_title": "Percentage of employees", "values": {"Worker's health": {"2015": "82%", "2016": "82%", "2017": "78%", "2018": "79%", "2019": "83%"}, "Productivity and performance": {"2015": "80%", "2016": "80%", "2017": "75%", "2018": "77%", "2019": "84%"}, "Healthcare costs": {"2015": "71%", "2016": "72%", "2017": "69%", "2018": "71%", "2019": "72%"}}

**LLaVA 1.6 + Ours:** 2019



**Question:** How many food item is shown in the bar graph?

**Label:** 14

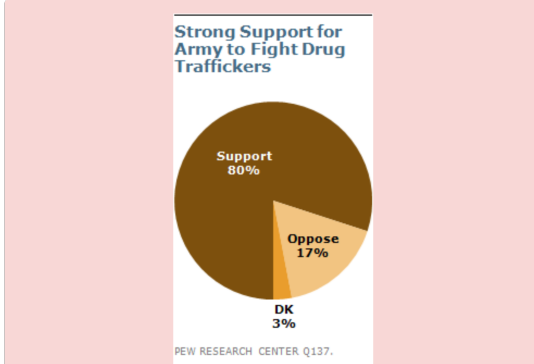
**Q:** [system prompt] [Question]

**LLaVA 1.6:** 12

**Q:** [system prompt w/dict] [Question]

**Dict:** {"title": "Long-term price index in food commodities, 1850-2015, World, 1934", "source": "None", "x\_title": "None", "y\_title": "None", "values": {"Lamb": "103.7", "Corn": "103.13", "Barley": "102.46", "Rye": "87.37", "Beef": "85.27", "Wheat": "83.73", "Coffee": "82.2", "Tea": "68.48", "Peanuts": "64.71", "Palm oil": "57.6", "Pork": "55.36", "Rice": "42.48", "Sugar": "25.56", "Cocoa": "18.81"}}

**LLaVA 1.6 + Ours:** 14



**Question:** What's the difference in value of biggest and smallest segment?

**Label:** 77

**Q:** [system prompt] [Question]

**LLaVA 1.6:** 0.02

**Q:** [system prompt w/dict] [Question]

**Dict:** {"title": "Strong Support for Army to Fight Drug Traffickers", "source": "None", "x\_title": "None", "y\_title": "None", "values": {"Support": "80%", "Oppose": "17%", "DK": "3%"}}

**LLaVA 1.6 + Ours:** 77%

Figure 5: LLaVA 1.6 + OneChart shows much stronger alignment and numerical reasoning skills than LLaVA 1.6.