

TUGAS UJIAN AKHIR SEMESTER (UAS)

EVALUASI PENDIDIKAN KEJURUAN

PENGEMBANGAN PENILAIAN PEMBELAJARAN

(BUTIR SOAL PILIHAN GANDA)

Disusun oleh:

ARIKA PRIHASTANTI SUTAMI

PROGRAM STUDI PENDIDIKAN GURU VOKASI

UNIVERSITAS SEBELAS MARET

2025

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga laporan tugas Ujian Akhir Semester (UAS) mata kuliah **Evaluasi Pendidikan Kejuruan Tahun 2025** ini dapat diselesaikan dengan baik. Laporan ini disusun sebagai bentuk pemenuhan tugas akademik sekaligus sebagai upaya pengembangan pemahaman penulis terhadap konsep, prinsip, dan praktik evaluasi pembelajaran, khususnya dalam pengembangan instrumen penilaian berbentuk soal pilihan ganda.

Penyusunan laporan ini difokuskan pada pengembangan penilaian pembelajaran yang meliputi penyusunan kisi-kisi soal, penulisan butir soal pilihan ganda, validasi isi menggunakan koefisien **Aiken's V**, serta analisis empiris butir soal menggunakan pendekatan **ITEMAN** berbasis *Classical Test Theory*. Melalui tahapan tersebut, diharapkan instrumen yang dikembangkan tidak hanya memenuhi kaidah teoretis evaluasi pendidikan, tetapi juga memiliki kualitas empiris yang baik ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

Penulis menyadari bahwa penyusunan instrumen penilaian yang berkualitas merupakan salah satu kompetensi penting bagi pendidik, khususnya dalam konteks pendidikan kejuruan yang menuntut pengukuran kemampuan secara objektif, adil, dan akurat. Oleh karena itu, laporan ini diharapkan dapat memberikan kontribusi akademik berupa pemahaman konseptual dan praktis mengenai pengembangan dan analisis soal pilihan ganda sebagai instrumen evaluasi pembelajaran.

Penulis mengucapkan terima kasih kepada dosen pengampu mata kuliah **Evaluasi Pendidikan Kejuruan** atas bimbingan, arahan, dan ilmu yang telah diberikan selama proses perkuliahan. Ucapan terima kasih juga disampaikan kepada semua pihak yang telah membantu secara langsung maupun tidak langsung dalam penyusunan laporan ini.

Penulis menyadari bahwa laporan ini masih memiliki keterbatasan dan memerlukan penyempurnaan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan demi perbaikan dan pengembangan keilmuan di masa yang akan datang. Semoga laporan ini dapat memberikan manfaat bagi pembaca dan menjadi referensi dalam pengembangan instrumen penilaian pembelajaran.

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
PENDAHULUAN	1
a. Teori Konseptual Soal Pilihan Ganda	1
METODE ANALISIS DATA	2
a. Aiken's V	2
b. Analisis Empiris Butir Soal Menggunakan Python	3
HASIL DAN PEMBAHASAN	4
a. Penyusunan Kisi-kisi	4
b. Butir Soal Pilihan Ganda dan Kunci Jawaban	6
c. Validasi Ahli dan Analisis Aiken's V	11
d. Lembar Validasi Ahli	12
e. Hasil Penghitungan Aiken's V dengan Python	23
1. Memindahkan data validasi ahli ke dalam excel (file bisa didownload dari github dibawah)	23
2. Membuat Coding Python untuk analisa Aiken's V	23
3. Hasil Analisa Aiken's V	25
4. Revisi Butir Soal Berdasarkan Masukan Ahli	26
5. Uji Coba Instrumen kepada Peserta Didik	26
6. Analisa Hasil Jawaban Siswa	27
7. Analisa Jawaban Siswa	28
a. Pembacaan Data	28
b. Skor Total Peserta	28
c. Tingkat Kesukaran	28
d. Daya Pembeda	28
e. Reliabilitas Tes (KR-20)	28
f. Analisis Distractor	28
8. Program Analisa	28

9. Hasil Analisa Jawaban Siswa	30
a Sheet Skor_Peserta	30
b Sheet Analisisi_Butir	30
1. Tingkat Kesukaran	30
2. Daya Pembeda	31
3. Keputusan Butir	31
c. Sheet Reliabilitas	31
d. Sheet Analisisi Distractor	32
KESIMPULAN	32
DAFTAR PUSTAKA	33

PENDAHULUAN

a. Teori Konseptual Soal Pilihan Ganda

Penilaian merupakan bagian integral dalam proses pembelajaran yang berfungsi untuk mengukur ketercapaian tujuan pembelajaran, menilai efektivitas proses pembelajaran, serta menjadi dasar pengambilan keputusan pendidikan. Melalui kegiatan penilaian, pendidik dapat memperoleh informasi yang akurat mengenai tingkat penguasaan kompetensi peserta didik, baik pada ranah kognitif, afektif, maupun psikomotor. Oleh karena itu, instrumen penilaian yang digunakan harus disusun secara sistematis dan memenuhi prinsip-prinsip pengukuran yang baik.

Salah satu instrumen penilaian yang paling banyak digunakan dalam praktik pendidikan formal adalah soal pilihan ganda. Bentuk soal ini termasuk ke dalam tes objektif yang memiliki keunggulan dalam hal objektivitas penskoran, efisiensi waktu, serta kemampuan menjangkau cakupan materi yang luas dalam waktu relatif singkat. Arikunto (2018) menyatakan bahwa tes objektif, termasuk soal pilihan ganda, memungkinkan penskoran dilakukan secara konsisten karena setiap butir soal telah dilengkapi dengan kunci jawaban yang pasti, sehingga meminimalkan subjektivitas penilai.

Selain keunggulan tersebut, soal pilihan ganda juga memiliki potensi untuk mengukur berbagai tingkat kemampuan kognitif, mulai dari kemampuan mengingat dan memahami hingga kemampuan menganalisis, mengevaluasi, dan menerapkan konsep, apabila dirancang dengan baik. Namun, dalam praktiknya, masih banyak soal pilihan ganda yang disusun tanpa memperhatikan kaidah penulisan dan prinsip pengukuran, sehingga kualitas soal yang dihasilkan belum optimal dan berpotensi menurunkan validitas hasil penilaian.

Soal pilihan ganda yang berkualitas tidak hanya ditentukan oleh bentuk dan jumlah alternatif jawaban, tetapi juga oleh proses pengembangan yang sistematis, kesesuaian dengan indikator kompetensi, serta hasil analisis empiris terhadap butir soal. Aspek-aspek penting seperti validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh merupakan indikator utama dalam menentukan kelayakan suatu butir soal untuk digunakan dalam penilaian.

Oleh karena itu, kajian teori mengenai pengembangan dan analisis soal pilihan ganda menjadi sangat penting sebagai dasar konseptual dalam evaluasi pendidikan. Landasan teori yang kuat tidak hanya membantu pendidik dalam menyusun instrumen penilaian yang bermutu, tetapi juga memastikan bahwa hasil penilaian yang diperoleh benar-benar mencerminkan

kemampuan peserta didik secara objektif, adil, dan dapat dipertanggungjawabkan secara ilmiah.

METODE ANALISIS DATA

Dalam penelitian pengembangan instrumen soal pilihan ganda, diperlukan prosedur analisis yang mampu menjamin kualitas instrumen baik dari sisi teoretis maupun empiris. Oleh karena itu, penelitian ini menggunakan **Aiken's V** dan **analisis statistik berbasis Python** sebagai dua pendekatan analisis yang saling melengkapi.

Aiken's V digunakan untuk mengukur **validitas isi** berdasarkan penilaian para ahli, sehingga setiap butir soal yang dikembangkan telah sesuai dengan indikator, tujuan pembelajaran, serta aspek konstruk dan bahasa. Selanjutnya, **bahasa pemrograman Python** digunakan untuk melakukan **analisis empiris butir soal** berdasarkan respons peserta didik, guna menilai tingkat kesukaran, daya pembeda, efektivitas pengecoh, dan reliabilitas tes. Penggunaan kedua teknik analisis ini diharapkan mampu menghasilkan instrumen soal pilihan ganda yang tidak hanya valid secara konseptual, tetapi juga andal dan efektif secara statistik.

a. Aiken's V

Aiken's V merupakan salah satu koefisien yang digunakan untuk mengukur **validitas isi (content validity)** suatu instrumen berdasarkan penilaian para ahli (*expert judgment*). Koefisien ini menunjukkan tingkat kesepakatan para ahli terhadap relevansi suatu butir instrumen dengan konstruk atau indikator yang hendak diukur. Aiken (1985) menyatakan bahwa validitas isi dapat diestimasi secara kuantitatif dengan menghitung proporsi kesepakatan para ahli terhadap setiap item.

Nilai Aiken's V berada pada rentang 0 sampai 1, di mana nilai yang mendekati 1 menunjukkan bahwa suatu butir soal dinilai sangat relevan oleh para ahli. Dalam penelitian ini, perhitungan Aiken's V dilakukan **menggunakan bahasa pemrograman Python** dengan memanfaatkan pustaka *pandas* untuk mengolah data penilaian ahli yang diperoleh dari lembar validasi. Hasil perhitungan digunakan untuk mengklasifikasikan butir soal ke dalam kategori relevansi tertentu sebagai dasar pengambilan keputusan revisi atau kelayakan butir soal sebelum uji empiris.

Validitas isi instrumen dianalisis menggunakan **koefisien Aiken's V**, berdasarkan penilaian dua orang ahli. Data penilaian diperoleh dari hasil validasi ahli yang diberikan pada skala 1–5.

Nilai Aiken's V dihitung menggunakan rumus:

$$V = \frac{\sum s}{n(c - 1)}$$

dengan:

- $s = r - l_o$
- r = skor yang diberikan validator
- l_o = skor terendah pada skala (1)
- n = jumlah validator
- c = jumlah kategori penilaian

Perhitungan dilakukan untuk setiap indikator soal menggunakan bantuan bahasa pemrograman **Python** dengan pustaka *pandas*.

Nilai Aiken's V yang diperoleh kemudian dikategorikan ke dalam lima tingkat relevansi, yaitu:

- $\leq 0,20$: Kurang Relevan
- $0,21-0,40$: Tidak Relevan
- $0,41-0,60$: Cukup Relevan
- $0,61-0,80$: Relevan
- $0,80$: Sangat Relevan

b. Analisis Empiris Butir Soal Menggunakan Python

Analisis empiris butir soal dalam penelitian ini dilakukan menggunakan **bahasa pemrograman Python** sebagai alternatif analisis setara ITEMAN yang berbasis **Classical Test Theory (CTT)**. Python digunakan untuk mengolah data respons peserta didik dan menghasilkan berbagai indikator statistik kualitas butir soal dan tes secara keseluruhan.

Analisis menggunakan Python meliputi:

1. Tingkat Kesukaran (Difficulty Index)

Menunjukkan proporsi peserta didik yang menjawab benar pada setiap butir soal. Butir soal dikategorikan mudah, sedang, atau sukar berdasarkan nilai indeks kesukaran yang diperoleh.

2. Daya Pembeda (Discrimination Index)

Menunjukkan kemampuan butir soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah. Butir dengan daya pembeda baik menunjukkan kualitas soal yang mampu mengukur kemampuan secara tepat.

3. Efektivitas Pengecoh (Distractor Effectiveness)

Menunjukkan sejauh mana pilihan jawaban salah berfungsi dengan baik. Pengecoh dikatakan efektif apabila dipilih oleh sebagian peserta didik, terutama dari kelompok berkemampuan rendah.

4. Reliabilitas Tes

Menggambarkan tingkat konsistensi internal instrumen secara keseluruhan. Nilai reliabilitas yang tinggi menunjukkan bahwa instrumen dapat dipercaya untuk mengukur kemampuan yang sama secara konsisten.

HASIL DAN PEMBAHASAN

a. Penyusunan Kisi-kisi

Tahap awal pengembangan instrumen diawali dengan penyusunan kisi-kisi soal sebagai pedoman dalam penulisan butir soal pilihan ganda. Kisi-kisi disusun berdasarkan indikator kompetensi dan tujuan pembelajaran yang telah ditetapkan, sehingga setiap butir soal memiliki keterkaitan yang jelas dengan aspek kemampuan yang diukur. Penyusunan kisi-kisi ini bertujuan untuk menjamin keterwakilan materi, kesesuaian level kognitif, serta keseimbangan distribusi soal.

Berikut kisi-kisinya:

No	Kompetensi yang Diuji	Level Kognitif	Indikator HOTS	Materi Pelajaran	Capaian Pembelajaran	No. Butir Soal
1	Memahami konsep dasar Python	C4 (Menganalisis)	Menafsirkan fungsi interpreter Python	Sintaks dasar Python, peran interpreter dalam eksekusi program	Memahami konsep dasar bahasa pemrograman Python dalam konteks analisis data	1–2
			Menganalisis peran interpreter dalam eksekusi program	Sintaks dasar Python, proses eksekusi kode	Menganalisis proses kerja Python sebagai dasar pengolahan data	3–4
2	Menginstal lingkungan kerja Python	C5 (Mengevaluasi)	Mengidentifikasi kesalahan instalasi Python	Instalasi Python, lingkungan pengembangan, library pendukung	Menerapkan lingkungan kerja Python untuk kebutuhan analisis data	5–6

No	Kompetensi yang Diuji	Level Kognitif	Indikator HOTS	Materi Pelajaran	Capaian Pembelajaran	No. Butir Soal
			Mengevaluasi kesesuaian sintaks dasar	Sintaks dasar Python, error sintaks	Mengidentifikasi kesalahan awal pemrograman Python	7–8
3	Menulis struktur dasar program Python	C4 (Menganalisis)	Menganalisis logika input output	Struktur program Python, input output data	Menulis kode Python secara sistematis untuk pengolahan data	9–10
			Mengkaji penggunaan variabel dalam program	Variabel, tipe data, struktur data sederhana	Menganalisis penggunaan variabel dalam pemrograman Python	11–12
4	Membuat program sederhana	C5 (Mengevaluasi)	Menilai kesalahan tipe data	Tipe data Python, struktur data	Mengevaluasi kesesuaian tipe data dalam pengolahan informasi	13–14
			Mengevaluasi penggunaan variabel dalam program	Variabel, struktur program sederhana	Mengintegrasikan konsep variabel dalam program Python	14–16
5	Sintesis proyek mini Python	C6 (Mencipta)	Merancang solusi logis berbasis studi kasus	Analisis data sederhana, penggunaan library Python	Mengintegrasikan konsep Python dalam pemecahan masalah berbasis data	17–18
			Mengembangkan algoritma sederhana kontekstual	Algoritma sederhana, konteks dunia nyata analisis data	Menghasilkan solusi berbasis data melalui proyek mini Python	19–20

b. Butir Soal Pilihan Ganda, Kunci Jawaban, dan Validasi Ahli

Kompetensi 1: Memahami Konsep Dasar Python

Indikator: Menafsirkan fungsi interpreter Python (No. 1–3)

1. Peran utama interpreter Python dalam menjalankan program adalah ...

- A. Mengubah kode Python menjadi file biner permanen
- B. Menerjemahkan kode baris per baris saat program dijalankan
- C. Mengoptimalkan kecepatan perangkat keras komputer
- D. Menyimpan hasil program ke dalam database
- E. Mengelola tampilan antarmuka pengguna

Kunci Jawaban: B

2. Jika sebuah baris kode Python mengandung kesalahan sintaks, interpreter akan ...

- A. Mengabaikan baris kode tersebut
- B. Tetap menjalankan seluruh program
- C. Menampilkan pesan kesalahan saat eksekusi
- D. Memperbaiki kesalahan secara otomatis
- E. Menghapus baris kode yang salah

Kunci Jawaban: C

Indikator: Menganalisis peran interpreter dalam eksekusi program (No. 4–5)

3. Seorang siswa menjalankan program Python dan mendapatkan error pada baris ke-5. Hal ini menunjukkan bahwa interpreter Python ...

- A. Mengeksekusi seluruh kode sebelum menemukan error
- B. Menjalankan kode secara bertahap sesuai urutan
- C. Mengabaikan error sintaks
- D. Menjalankan program secara paralel
- E. Menyimpan kode ke dalam cache

Kunci Jawaban: B

4. Dibandingkan bahasa compiled, penggunaan interpreter Python berdampak pada ...

- A. Hilangnya fleksibilitas debugging
- B. Ketergantungan pada compiler
- C. Kemudahan pengujian kode secara langsung

- D. Tidak adanya error sintaks
- E. Kecepatan eksekusi selalu lebih tinggi

Kunci Jawaban: C

Kompetensi 2: Menginstal Lingkungan Kerja Python

Indikator: Mengidentifikasi kesalahan instalasi Python (No. 6–8)

5. Salah satu indikasi Python belum terinstal dengan benar pada sistem adalah ...

- A. Program Python berjalan lambat
- B. Perintah python --version tidak dikenali
- C. File .py tidak dapat disimpan
- D. IDE tidak dapat dibuka
- E. Program menghasilkan output kosong

Kunci Jawaban: B

6. Kesalahan paling umum saat instalasi Python pada Windows adalah ...

- A. Tidak menginstal library tambahan
- B. Tidak mencentang opsi *Add Python to PATH*
- C. Menggunakan editor teks
- D. Menjalankan Python melalui terminal
- E. Menggunakan sistem operasi 64-bit

Kunci Jawaban: B

Indikator: Mengevaluasi kesesuaian sintaks dasar (No. 9–10)

7. Pernyataan berikut yang menyebabkan error sintaks adalah ...

- A. `print("Hello World")`
- B. `x = 10`
- C. `if x > 5:`
- D. `print "Hello"`
- E. `y = x + 2`

Kunci Jawaban: D

8. Kesalahan penulisan indentasi pada Python akan berdampak pada ...

- A. Kesalahan logika program
- B. Error saat proses eksekusi

- C. Tampilan output tidak rapi
- D. Ukuran file program
- E. Kecepatan program

Kunci Jawaban: B

Kompetensi 3: Menulis Struktur Dasar Program Python

Indikator: Menganalisis logika input output (No. 11–13)

9. Fungsi utama perintah input() dalam Python adalah ...

- A. Menampilkan output ke layar
- B. Mengambil data dari file
- C. Menerima masukan dari pengguna
- D. Menyimpan data ke variabel
- E. Mengonversi tipe data

Kunci Jawaban: C

10.

```
x = int(input())
```

```
print(x * 2)
```

Jika pengguna memasukkan angka 5, output program adalah ...

- A. 5
- B. 7
- C. 10
- D. 25
- E. Error

Kunci Jawaban: C

Indikator: Mengkaji penggunaan variabel dalam program (No. 14–15)

11. Variabel dalam Python digunakan untuk ...

- A. Menyimpan perintah program
- B. Menyimpan data sementara
- C. Menyimpan hasil instalasi
- D. Menyimpan error program
- E. Menjalankan perintah

Kunci Jawaban: B

12. Penamaan variabel berikut yang sesuai dengan aturan Python adalah ...

- A. 2data
- B. data nilai
- C. data_nilai
- D. data-nilai
- E. class

Kunci Jawaban: C

Kompetensi 4: Membuat Program Sederhana

Indikator: Menilai kesalahan tipe data (No. 16–18)

13. Kesalahan tipe data terjadi pada pernyataan ...

- A. $x = 10$
- B. $y = "5"$
- C. $z = x + y$
- D. `print(x)`
- E. $a = 3.5$

Kunci Jawaban: C

14. Untuk menghindari kesalahan tipe data saat operasi aritmetika, langkah yang tepat adalah ...

- A. Menghapus variabel
- B. Menggunakan komentar
- C. Melakukan konversi tipe data
- D. Menjalankan ulang program
- E. Mengganti editor

Kunci Jawaban: C

Indikator: Mengevaluasi penggunaan variabel dalam program (No. 19–20)

15. Penggunaan variabel yang tidak tepat dapat menyebabkan ...

- A. Program lebih cepat
- B. Program lebih efisien
- C. Kesalahan logika program

- D. Output selalu benar
- E. Program otomatis berhenti

Kunci Jawaban: C

16. Evaluasi yang benar terhadap variabel dalam program Python adalah ...

- A. Variabel boleh digunakan tanpa didefinisikan
- B. Variabel harus konsisten dengan tipe datanya
- C. Variabel hanya digunakan satu kali
- D. Variabel tidak boleh diubah nilainya
- E. Variabel tidak memengaruhi output

Kunci Jawaban: B

Kompetensi 5: Sintesis Proyek Mini Python

Indikator: Merancang solusi logis berbasis studi kasus (No. 21–23)

17. Dalam proyek analisis data sederhana, langkah pertama yang logis adalah ...

- A. Menampilkan grafik
- B. Menghapus data
- C. Mengidentifikasi permasalahan
- D. Menggunakan library Matplotlib
- E. Menyimpan hasil ke file

Kunci Jawaban: C

18. Jika tujuan proyek adalah menganalisis nilai siswa, data yang paling relevan adalah ...

- A. Warna grafik
- B. Nama file program
- C. Nilai numerik siswa
- D. Versi Python
- E. Editor yang digunakan

Kunci Jawaban: C

Indikator: Mengembangkan algoritma sederhana kontekstual (No. 24–25)

19. Algoritma sederhana dalam Python ditandai dengan ...

- A. Banyaknya baris kode
- B. Alur logika yang jelas dan terstruktur

- C. Penggunaan library kompleks
- D. Tidak menggunakan variabel
- E. Tidak menghasilkan output

Kunci Jawaban: B

20. Dalam konteks dunia nyata, penggunaan Python untuk analisis data bertujuan untuk ...

- A. Menggantikan semua bahasa pemrograman
- B. Menampilkan kode yang kompleks
- C. Mendukung pengambilan keputusan berbasis data
- D. Mengurangi kebutuhan data
- E. Menghilangkan proses analisis

Kunci Jawaban: C

c. Validasi Ahli dan Analisis Aiken's V

Validasi ahli dilakukan untuk menilai **validitas isi** instrumen soal pilihan ganda sebelum dilakukan uji empiris. Penilaian dilakukan oleh ahli evaluasi pembelajaran dan ahli materi Python terhadap aspek **kesesuaian indikator, konstruk soal, dan penggunaan bahasa**, menggunakan skala Likert 1–5. Data hasil penilaian para ahli selanjutnya dianalisis menggunakan **koefisien Aiken's V yang dihitung dengan bantuan bahasa pemrograman Python**, untuk mengetahui tingkat kesepakatan ahli terhadap relevansi setiap butir soal.

Nilai Aiken's V berada pada rentang 0 sampai 1, di mana nilai yang mendekati 1 menunjukkan bahwa suatu butir soal memiliki **validitas isi yang tinggi**.

Berikut indikator penilaian validasi ahli instrumen soal pilihan ganda:

1. Kesesuaian Isi

- Kesesuaian butir soal dengan indikator kompetensi dan capaian pembelajaran Python.
- Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).

2. Konstruk Soal

- Kesesuaian perumusan pokok soal (stem) yang jelas, fokus, dan tidak memberikan petunjuk ke arah jawaban benar.
- Kesesuaian soal dengan keberadaan satu jawaban yang paling benar serta alternatif jawaban yang homogen, logis, setara, dan distractor berfungsi secara konseptual.

3. Bahasa

- Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik dan mudah dipahami.
- Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, dan tidak menimbulkan ambiguitas makna.

d. Lembar Validasi Ahli

LEMBAR VALIDASI AHLI

A. Identitas Validator

Nama Validator :

Keahlian : ☐ Pendidikan ☐ Pemrograman ☐ Evaluasi Pembelajaran
☐ Lainnya:

Instansi :

Jabatan :

Tanggal Penilaian :

B. Petunjuk Pengisian

Berilah tanda centang (✓) pada kolom yang sesuai dengan penilaian Bapak/Ibu terhadap setiap pernyataan pada masing-masing soal.

Keterangan Skala Penilaian:

- 1 = Sangat Tidak Relevan
2 = Tidak Relevan
3 = Cukup Relevan
4 = Relevan
5 = Sangat Relevan

C. Instrumen Validasi Ahli

Pilihlah jawaban dari pertanyaan di bawah ini!

Kompetensi 1: Memahami Konsep Dasar Python

Indikator: Menafsirkan fungsi interpreter Python (No. 1–3)

1. Peran utama interpreter Python dalam menjalankan program adalah ...
- A. Mengubah kode Python menjadi file biner permanen
- B. Menerjemahkan kode baris per baris saat program dijalankan
- C. Mengoptimalkan kecepatan perangkat keras komputer

D. Menyimpan hasil program ke dalam database

E. Mengelola tampilan antarmuka pengguna

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

2. Jika sebuah baris kode Python mengandung kesalahan sintaks, interpreter akan ...

A. Mengabaikan baris kode tersebut

B. Tetap menjalankan seluruh program

C. Menampilkan pesan kesalahan saat eksekusi

D. Memperbaiki kesalahan secara otomatis

E. Menghapus baris kode yang salah

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Indikator: Menganalisis peran interpreter dalam eksekusi program (No. 4–5)

3. Seorang siswa menjalankan program Python dan mendapatkan error pada baris ke-5. Hal ini menunjukkan bahwa interpreter Python ...

A. Mengeksekusi seluruh kode sebelum menemukan error

B. Menjalankan kode secara bertahap sesuai urutan

C. Mengabaikan error sintaks

D. Menjalankan program secara paralel

E. Menyimpan kode ke dalam cache

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

4. Dibandingkan bahasa compiled, penggunaan interpreter Python berdampak pada ...

A. Hilangnya fleksibilitas debugging

B. Ketergantungan pada compiler

C. Kemudahan pengujian kode secara langsung

D. Tidak adanya error sintaks

E. Kecepatan eksekusi selalu lebih tinggi

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Kompetensi 2: Menginstal Lingkungan Kerja Python

Indikator: Mengidentifikasi kesalahan instalasi Python (No. 6–8)

5. Salah satu indikasi Python belum terinstal dengan benar pada sistem adalah ...

A. Program Python berjalan lambat

B. Perintah python --version tidak dikenali

C. File .py tidak dapat disimpan

D. IDE tidak dapat dibuka

E. Program menghasilkan output kosong

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

6. Kesalahan paling umum saat instalasi Python pada Windows adalah ...

A. Tidak menginstal library tambahan

B. Tidak mencentang opsi *Add Python to PATH*

C. Menggunakan editor teks

D. Menjalankan Python melalui terminal

E. Menggunakan sistem operasi 64-bit

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Indikator: Mengevaluasi kesesuaian sintaks dasar (No. 9–10)

7. Pernyataan berikut yang menyebabkan error sintaks adalah ...

A. `print("Hello World")`

B. `x = 10`

C. `if x > 5:`

D. print "Hello"

E. $y = x + 2$

Kunci Jawaban: D

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

8. Kesalahan penulisan indentasi pada Python akan berdampak pada ...

A. Kesalahan logika program

B. Error saat proses eksekusi

C. Tampilan output tidak rapi

D. Ukuran file program

E. Kecepatan program

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Kompetensi 3: Menulis Struktur Dasar Program Python

Indikator: Menganalisis logika input output (No. 11–13)

9. Fungsi utama perintah input() dalam Python adalah ...

A. Menampilkan output ke layar

B. Mengambil data dari file

C. Menerima masukan dari pengguna

D. Menyimpan data ke variabel

E. Mengonversi tipe data

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

10.

```
x = int(input())
```

```
print(x * 2)
```

Jika pengguna memasukkan angka 5, output program adalah ...

A. 5

B. 7

C. 10

D. 25

E. Error

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Indikator: Mengkaji penggunaan variabel dalam program (No. 14–15)

11. Variabel dalam Python digunakan untuk ...

A. Menyimpan perintah program

- B. Menyimpan data sementara
- C. Menyimpan hasil instalasi
- D. Menyimpan error program
- E. Menjalankan perintah

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

12. Penamaan variabel berikut yang sesuai dengan aturan Python adalah ...

- A. 2data
- B. data nilai
- C. data_nilai
- D. data-nilai
- E. class

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Kompetensi 4: Membuat Program Sederhana

Indikator: Menilai kesalahan tipe data (No. 16–18)

13. Kesalahan tipe data terjadi pada pernyataan ...

- A. $x = 10$

- B. $y = "5"$
 C. $z = x + y$
 D. `print(x)`
 E. $a = 3.5$

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

14. Untuk menghindari kesalahan tipe data saat operasi aritmetika, langkah yang tepat adalah ...

- A. Menghapus variabel
 B. Menggunakan komentar
 C. Melakukan konversi tipe data
 D. Menjalankan ulang program
 E. Mengganti editor

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Indikator: Mengevaluasi penggunaan variabel dalam program (No. 19–20)

15. Penggunaan variabel yang tidak tepat dapat menyebabkan ...

- A. Program lebih cepat

- B. Program lebih efisien
- C. Kesalahan logika program
- D. Output selalu benar
- E. Program otomatis berhenti

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

16. Evaluasi yang benar terhadap variabel dalam program Python adalah ...

- A. Variabel boleh digunakan tanpa didefinisikan
- B. Variabel harus konsisten dengan tipe datanya
- C. Variabel hanya digunakan satu kali
- D. Variabel tidak boleh diubah nilainya
- E. Variabel tidak memengaruhi output

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Kompetensi 5: Sintesis Proyek Mini Python

Indikator: Merancang solusi logis berbasis studi kasus (No. 21–23)

17. Dalam proyek analisis data sederhana, langkah pertama yang logis adalah ...

- A. Menampilkan grafik

- B. Menghapus data
- C. Mengidentifikasi permasalahan
- D. Menggunakan library Matplotlib
- E. Menyimpan hasil ke file

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

18. Jika tujuan proyek adalah menganalisis nilai siswa, data yang paling relevan adalah ...

- A. Warna grafik
- B. Nama file program
- C. Nilai numerik siswa
- D. Versi Python
- E. Editor yang digunakan

Kunci Jawaban: C

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

Indikator: Mengembangkan algoritma sederhana kontekstual (No. 24–25)

19. Algoritma sederhana dalam Python ditandai dengan ...

- A. Banyaknya baris kode
- B. Alur logika yang jelas dan terstruktur

- C. Penggunaan library kompleks
- D. Tidak menggunakan variabel
- E. Tidak menghasilkan output

Kunci Jawaban: B

No	Item Pernyataan	Skala Nilai				
		1	2	3	4	5
1	Kesesuaian butir soal dengan indikator kompetensi serta capaian pembelajaran Python.					
2	Kesesuaian substansi soal dengan konteks pembelajaran Python serta tuntutan kemampuan berpikir tingkat tinggi (HOTS).					
3	Kesesuaian perumusan pokok soal yang jelas, fokus, serta tidak memberikan petunjuk ke arah jawaban benar.					
4	Kesesuaian soal dengan keberadaan satu jawaban paling benar serta alternatif jawaban yang homogen, logis, setara, serta distractor berfungsi secara konseptual.					
5	Kesesuaian penggunaan bahasa dalam soal dengan kaidah bahasa Indonesia yang baik serta mudah dipahami.					
6	Kesesuaian penggunaan istilah pemrograman Python secara tepat, konsisten, serta tidak menimbulkan ambiguitas makna.					

20. Dalam konteks dunia nyata, penggunaan Python untuk analisis data bertujuan untuk ...

- A. Menggantikan semua bahasa pemrograman
- B. Menampilkan kode yang kompleks
- C. Mendukung pengambilan keputusan berbasis data
- D. Mengurangi kebutuhan data
- E. Menghilangkan proses analisis

Kunci Jawaban: C

H. Saran dan Catatan Validator

.....

.....

.....

.....,

Validator,

(_____)

e. Hasil Penghitungan Aiken's V dengan Python

Penghitungan koefisien Aiken's V dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan data hasil penilaian para ahli. Skor penilaian dianalisis untuk mengetahui tingkat kesepakatan ahli terhadap relevansi setiap butir soal. Hasil perhitungan Aiken's V digunakan sebagai dasar penentuan kelayakan butir soal, apakah dinyatakan layak, perlu direvisi, atau dieliminasi sebelum dilakukan uji empiris.

Langkah analisa:

1. Memindahkan data validasi ahli ke dalam excel (file bisa didownload dari github dibawah)

https://raw.githubusercontent.com/arikaps/analisissoal/main/data_validasi_siappython.xlsx

Berikut sebagai preview dari data lengkap sejumlah 120 baris:

	A	B	C	D	E
1	No_Soal	Aspek	Indikator	Skor_V1	Skor_V2
2		1 Kesesuaian	Sesuai indikator & CP	4	5
3		1 Kesesuaian	Relevansi & HOTS	5	5
4		1 Konstruk	Kejelasan rumusan soal	5	4
5		1 Konstruk	Jawaban benar & distraktor	4	4
6		1 Bahasa	Kejelasan bahasa	5	5
7		1 Bahasa	Konsistensi istilah Python	5	4
8		2 Kesesuaian	Sesuai indikator & CP	5	5
9		2 Kesesuaian	Relevansi & HOTS	5	5

2. Membuat Coding Python untuk analisa Aiken's V

Coding Python lengkap bisa didownload di github dengan link:

https://raw.githubusercontent.com/arikaps/analisissoal/main/analisis_aiken_v.py

Ringkasan Coding:

```
import pandas as pd

# Membaca data hasil validasi ahli
file_path =
"https://raw.githubusercontent.com/arikaps/analisissoal/main/data_validasi_siappython.xlsx"

df = pd.read_excel(file_path)

# Parameter Aiken's V
lo = 1          # skor terendah skala
c = 5           # jumlah kategori penilaian
validator_cols = ["Skor_V1", "Skor_V2"]
```

```

n = len(validator_cols)

# Menghitung Aiken's V
s = df[validator_cols] - lo
df["Aiken_V"] = s.sum(axis=1) / (n * (c - 1))

# Kategori relevansi
def kategori_relevansi(v):
    if v <= 0.20:
        return "Kurang Relevan"
    elif v <= 0.40:
        return "Tidak Relevan"
    elif v <= 0.60:
        return "Cukup Relevan"
    elif v <= 0.80:
        return "Relevan"
    else:
        return "Sangat Relevan"

df["Kategori_Relevansi"] = df["Aiken_V"].apply(kategori_relevansi)

# Rata-rata Aiken's V per soal
aiken_per_soal = (
    df.groupby("No_Soal")["Aiken_V"]
    .mean()
    .reset_index()
)

```

Kode Python ini digunakan untuk **menghitung validitas isi butir soal menggunakan koefisien Aiken's V** berdasarkan penilaian dua orang ahli. Data validasi dibaca langsung dari file Excel yang tersimpan di GitHub menggunakan pustaka *pandas*. Selanjutnya ditetapkan parameter skala penilaian (skor terendah, jumlah kategori, dan jumlah validator), kemudian nilai Aiken's V dihitung untuk setiap butir soal dengan mengonversi skor ahli ke dalam nilai $s = r - l_o$. Nilai Aiken's V yang diperoleh diklasifikasikan ke dalam kategori relevansi, mulai dari kurang relevan hingga sangat relevan. Terakhir, dilakukan perhitungan rata-rata nilai Aiken's V untuk setiap nomor soal guna memperoleh gambaran tingkat validitas isi soal secara keseluruhan.

3. Hasil Analisa Aiken's V

Hasil analisis validitas isi menggunakan koefisien Aiken's V menunjukkan bahwa nilai Aiken's V setiap indikator berada pada rentang **0,75 hingga 1,00**. Sebagian besar indikator memperoleh nilai **Aiken's V $\geq 0,80$** dan termasuk dalam kategori **sangat relevan**, sedangkan sebagian kecil indikator berada pada kategori **relevan** dengan nilai Aiken's V sebesar **0,75**. Tidak ditemukan indikator dengan kategori cukup relevan, tidak relevan, maupun kurang relevan.

Rata-rata nilai Aiken's V per soal berada pada kisaran **0,89–0,96**, yang menunjukkan tingkat kesepakatan ahli yang tinggi terhadap kesesuaian isi, konstruk, dan bahasa pada butir soal yang dikembangkan. Dengan demikian, secara keseluruhan instrumen soal dapat dinyatakan **memiliki validitas isi yang sangat baik** dan **layak digunakan** tanpa revisi substansial, meskipun beberapa indikator dengan nilai lebih rendah masih dapat disempurnakan secara minor.

Hasil analisa Aiken's V dengan menggunakan Python berupa output yang sudah diupload ke github dengan link:

https://raw.githubusercontent.com/arikaps/analisissoal/main/hasil_aiken_v_relevansi.xlsx

No_Soal	Aspek	Indikator	Skor_V1	Skor_V2	s_Skor_V1	s_Skor_V2	Aiken_V	gori_Relevansi
1	Kesesuaia	Sesuai ind	4	5	3	4	0.875	Sangat Relevan
1	Kesesuaia	Relevansi	5	5	4	4	1	Sangat Relevan
1	Konstruk	Kejelasan	5	4	4	3	0.875	Sangat Relevan
1	Konstruk	Jawaban k	4	4	3	3	0.75	Relevan
1	Bahasa	Kejelasan	5	5	4	4	1	Sangat Relevan
1	Bahasa	Konsisten	5	4	4	3	0.875	Sangat Relevan
2	Kesesuaia	Sesuai ind	5	5	4	4	1	Sangat Relevan
2	Kesesuaia	Relevansi	5	5	4	4	1	Sangat Relevan
2	Konstruk	Kejelasan	5	5	4	4	1	Sangat Relevan
2	Konstruk	Jawaban k	4	5	3	4	0.875	Sangat Relevan
2	Bahasa	Kejelasan	4	4	3	3	0.75	Relevan
2	Bahasa	Konsisten	4	4	3	3	0.75	Relevan
3	Kesesuaia	Sesuai ind	5	5	4	4	1	Sangat Relevan

Aiken_per_Indikator		Aiken_per_Soal		+
No_Soal	Aiken_V_Rata2			
1	0.895833			
2	0.895833			
3	0.895833			
4	0.958333			
5	0.916667			
6	0.895833			
7	0.9375			
8	0.875			
9	0.9375			
10	0.958333			
11	1			
12	0.895833			
13	0.895833			

4. Revisi Butir Soal Berdasarkan Masukan Ahli

Berdasarkan hasil penilaian ahli yang dianalisis menggunakan koefisien Aiken's V, secara umum butir-butir soal yang dikembangkan berada pada kategori **relevan hingga sangat relevan**, sehingga instrumen dinyatakan memiliki validitas isi yang baik dan layak digunakan. Oleh karena itu, **tidak diperlukan revisi substansial** terhadap butir soal. Namun demikian, beberapa masukan dari ahli tetap dijadikan bahan pertimbangan untuk **penyempurnaan redaksi dan kejelasan butir soal**.

Revisi yang dilakukan bersifat minor, meliputi perbaikan penggunaan bahasa agar lebih jelas, penyederhanaan kalimat agar tidak menimbulkan penafsiran ganda, serta penyesuaian istilah agar lebih sesuai dengan konteks materi yang diukur. Dengan adanya revisi tersebut, diharapkan butir soal menjadi lebih mudah dipahami oleh responden tanpa mengubah substansi atau indikator yang diukur, sehingga kualitas instrumen penelitian semakin optimal.

5. Uji Coba Instrumen kepada Peserta Didik

Setelah instrumen dinyatakan memiliki validitas isi yang memadai berdasarkan penilaian ahli dan analisis Aiken's V, langkah selanjutnya adalah **melakukan uji coba instrumen kepada peserta didik**. Uji coba ini bertujuan untuk memperoleh data empirik mengenai kualitas butir soal saat digunakan oleh responden sesungguhnya.

Instrumen diberikan kepada peserta didik yang memiliki karakteristik serupa dengan subjek penelitian. Pelaksanaan uji coba dilakukan dalam kondisi yang terkendali sesuai dengan prosedur pelaksanaan tes. Data hasil uji coba berupa jawaban peserta didik selanjutnya dikumpulkan dan disiapkan untuk dianalisis pada tahap berikutnya.

Instrumen yang diberikan kepada siswa berupa soal, sedangkan jawaban siswa akan diuji untuk tingkat kesukaran (difficulty index), daya pembeda (discrimination index), efektivitas pengecoh (distractor effectiveness), dan reliabilitas tes. Berikut preview hasil dari Uji Coba dari siswa sejumlah 60 siswa:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	ID_Peserta	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
2	P1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	0
3	P2	1	1	1	0	0	1	1	1	0	1	1	0	1	1	1
4	P3	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1
5	P4	1	0	0	0	1	1	0	1	1	1	0	1	0	1	0
6	P5	0	1	0	0	1	1	0	1	0	1	1	1	0	0	1
7	P6	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1
8	P7	0	1	0	1	0	0	1	1	0	0	1	1	1	0	0
9	P8	1	1	1	0	0	0	0	1	0	1	1	0	1	1	1
10	P9	0	1	0	0	1	0	0	1	0	1	0	0	1	1	0
11	P10	1	1	0	1	0	1	1	1	1	1	0	0	1	0	1
12	P11	1	1	1	1	1	0	1	0	0	0	1	0	0	1	1
13	P12	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1
14	P13	1	1	0	1	1	0	1	1	0	0	0	1	1	1	1

Nilai 1 berarti jawaban benar dan 0 bernilai salah dalam memilih jawaban dari soal. Untuk hasil excel dapat di download dari link:

https://raw.githubusercontent.com/arikaps/analisoal/main/data_jawaban_siswa.xlsx

6. Analisa Hasil Jawaban Siswa

Analisis hasil jawaban siswa dilakukan untuk mengetahui pola pencapaian peserta terhadap butir soal yang diberikan. Data jawaban siswa dianalisis berdasarkan skor benar–salah, distribusi pilihan jawaban, serta keterkaitannya dengan tingkat kesukaran, daya pembeda, dan efektivitas distractor. Berikut preview hasil jawaban siswa yang terangkum ke dalam excel:

- Sheet nilai

ID_Peserta	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
P1	1	1	1	1	1	0	1	0	0	1	0	1
P2	1	1	1	0	0	1	1	1	0	1	1	0
P3	0	1	0	1	1	0	0	0	0	0	0	0
P4	1	0	0	0	1	1	0	1	1	1	0	1
P5	0	1	0	0	1	1	0	1	0	1	1	1
P6	1	0	1	1	1	1	0	1	0	1	0	1
P7	0	1	0	1	0	0	1	1	0	0	1	1
P8	1	1	1	0	0	0	0	1	0	1	1	0

- Sheet opsi

S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
B	C	B	C	B	C	D	E	E	C	C	C
B	C	B	B	A	B	D	B	B	C	B	B
E	C	E	C	B	C	C	E	A	D	E	B
B	A	A	E	B	B	C	B	C	C	A	C
E	C	E	E	B	B	A	B	E	C	B	C
B	E	B	C	B	B	A	B	E	C	E	C
E	C	E	C	C	A	D	B	D	E	B	C
B	C	B	D	E	E	E	B	A	C	B	B
C	C	E	A	B	A	B	B	E	C	D	A
B	C	C	C	A	B	D	B	C	C	A	E

- Sheet kunci

no_soal	kunci
1	B
2	C
3	B
4	C
5	B
6	B

Untuk file lengkap dapat di download di github dengan link:

https://raw.githubusercontent.com/arikaps/analisoal/main/jawaban_siswa_nilai_opsi_kunci.xlsx

7. Analisa Jawaban Siswa

Seluruh jawaban siswa dianalisa dengan menggunakan Python. Bebeapa analisa diantaranya:

a. Pembacaan Data

Data dianalisis menggunakan tiga lembar kerja Excel, yaitu nilai (skor benar–salah), opsi (jawaban pilihan ganda), dan kunci (jawaban benar). Pemisahan ini bertujuan menjaga konsistensi data dan memudahkan analisis lanjutan.

b. Skor Total Peserta

Skor total diperoleh dengan menjumlahkan skor jawaban benar pada seluruh butir soal untuk setiap peserta. Skor ini digunakan sebagai dasar pengelompokan peserta pada analisis daya pembeda dan perhitungan reliabilitas.

c. Tingkat Kesukaran

Tingkat kesukaran dihitung sebagai rata-rata skor jawaban benar pada setiap butir soal. Nilai ini menunjukkan proporsi peserta yang mampu menjawab soal dengan benar dan digunakan untuk mengklasifikasikan butir menjadi soal sukar, sedang, atau mudah.

d. Daya Pembeda

Daya pembeda dihitung menggunakan metode kelompok atas dan bawah 27%. Selisih proporsi jawaban benar antara kedua kelompok menunjukkan kemampuan butir soal dalam membedakan peserta berkemampuan tinggi dan rendah.

e. Reliabilitas Tes (KR-20)

Reliabilitas tes dihitung menggunakan rumus Kuder–Richardson 20 (KR-20) yang sesuai untuk instrumen dengan skor dikotomi. Nilai reliabilitas menunjukkan tingkat konsistensi internal instrumen secara keseluruhan.

f. Analisis Distractor

Analisis distractor dilakukan dengan menghitung proporsi pemilihan setiap opsi pengecoh (selain kunci). Distractor dinyatakan berfungsi apabila dipilih oleh minimal 5% peserta, yang menunjukkan bahwa opsi tersebut mampu mengecoh peserta yang belum menguasai materi.

8. Program Analisa

Berikut Preview Code Programnya:

```
import pandas as pd
import numpy as np

# =====
# 1. Membaca data (3 sheet)
```

```

# =====
file_path = "jawaban_siswa_nilai_opsi_kunci.xlsx"

nilai = pd.read_excel(file_path, sheet_name="nilai")
opsi = pd.read_excel(file_path, sheet_name="opsi")
kunci = pd.read_excel(file_path, sheet_name="kunci")

item = [c for c in nilai.columns if c.startswith("S")]

# =====
# 2. Skor total
# =====
nilai["Skor_Total"] = nilai[item].sum(axis=1)

# =====
# 3. Tingkat kesukaran
# =====
p = nilai[item].mean()

# =====
# 4. Daya pembeda (27%)
# =====
nilai_urut = nilai.sort_values("Skor_Total", ascending=False)
n27 = int(0.27 * len(nilai))

D = nilai_urut.head(n27)[item].mean() -
nilai_urut.tail(n27)[item].mean()

# =====
# 5. Reliabilitas KR-20
# =====
q = 1 - p
var_total = nilai["Skor_Total"].var(ddof=1)

kr20 = (len(item) / (len(item) - 1)) * (1 - (np.sum(p * q) /
var_total))

# =====
# 6. Analisis distractor
# =====
kunci_dict = dict(zip("S" + kunci["no_soal"].astype(str),
kunci["kunci"]))

hasil_distractor = []

for s in item:
    distribusi = opsi[s].value_counts()
    for o in ["A", "B", "C", "D", "E"]:
        if o != kunci_dict[s]:
            prop = distribusi.get(o, 0) / distribusi.sum()
            hasil_distractor.append([s, o, round(prop, 3)])

```

Untuk keseluruhan kode program bisa didownload di github:

https://raw.githubusercontent.com/arikaps/analisisoal/main/analisis_butir_soal.py

9. Hasil Analisa Jawaban Siswa

Berikut Preview hasil analisa jawaban siswa:

a Sheet Skor_Peserta

	A	B	C
1	ID_Peserta	Skor_Total	
2	P1	13	
3	P2	15	
4	P3	9	
5	P4	12	
6	P5	12	
7	P6	13	
8	P7	11	
9	P8	11	
10	P9	8	
11	P10	13	
12	P11	14	
13	P12	16	
14	P13	12	

Sheet Skor_Peserta menyajikan hasil perolehan skor total setiap peserta tes. Skor total diperoleh dari penjumlahan jawaban benar pada seluruh butir soal (skor 1 untuk jawaban benar dan 0 untuk jawaban salah). Data pada sheet ini digunakan sebagai dasar dalam pengelompokan peserta ke dalam kelompok atas dan kelompok bawah pada analisis daya pembeda, serta menjadi dasar perhitungan reliabilitas tes.

b Sheet Analisisi_Butir

Butir_Soal	Kat_Kesuk	gori_Kesuk	ya_Pemb	gori_Daya	Keputusan
S1	0.555556	Sedang	0.9	Sangat Baik	Diterima
S2	0.75	Mudah	0.05	Jelek	Dibuang
S3	0.555556	Sedang	0.1	Jelek	Dibuang
S4	0.555556	Sedang	0.325	Baik	Diterima
S5	0.666667	Sedang	-0.05	Jelek	Dibuang
S6	0.638889	Sedang	0.3	Baik	Diterima
S7	0.555556	Sedang	0.225	Cukup	Direvisi
S8	0.611111	Sedang	0.45	Sangat Baik	Diterima

1. Tingkat Kesukaran

- Hampir seluruh butir berada pada kategori sedang ($p \approx 0,44-0,69$), sehingga secara umum tes tidak terlalu mudah maupun terlalu sulit.
- Hanya S2 yang tergolong mudah ($p = 0,75$).

2. Daya Pembeda

- Terdapat butir dengan daya pembeda sangat baik (misalnya S1, S8, S11, S16, S18, S19) yang mampu membedakan peserta berkemampuan tinggi dan rendah dengan baik.
- Beberapa butir memiliki daya pembeda cukup (S7, S12) sehingga masih dapat digunakan setelah perbaikan.
- Banyak butir menunjukkan daya pembeda jelek atau negatif (misalnya S2, S3, S5, S10, S13, S14, S15, S17, S20) yang menandakan butir tidak efektif membedakan kemampuan peserta.

3. Keputusan Butir

Diterima: 9 butir (S1, S4, S6, S8, S9, S11, S16, S18, S19) karena memiliki kesukaran sesuai dan daya pembeda baik–sangat baik.

- Direvisi: 2 butir (S7, S12) karena daya pembeda masih cukup.
- Dibuang: 9 butir karena daya pembeda rendah atau negatif meskipun tingkat kesukarannya sedang.

Kesimpulan

Instrumen memiliki tingkat kesukaran yang relatif seimbang, namun kualitas butir bervariasi. Sekitar setengah butir sudah baik dan layak digunakan, sedangkan sisanya perlu direvisi atau dibuang untuk meningkatkan kualitas dan reliabilitas tes secara keseluruhan.

c. Sheet Reliabilitas

is_Reliabili	Reliabil	Interpretasi
KR-20	0.607	Sedang

Nilai reliabilitas instrumen berdasarkan koefisien **KR-20 sebesar 0,607** menunjukkan bahwa tes memiliki **tingkat reliabilitas sedang**. Artinya, hasil tes sudah cukup konsisten dalam mengukur kemampuan peserta, namun masih terdapat peluang terjadinya kesalahan pengukuran. Dengan reliabilitas kategori sedang, instrumen dapat digunakan untuk keperluan evaluasi pembelajaran, tetapi akan lebih baik jika dilakukan penyempurnaan, misalnya dengan merevisi atau mengeliminasi butir soal yang memiliki daya pembeda rendah atau distractor yang tidak berfungsi, agar reliabilitas tes dapat meningkat.

d. Sheet Analisis Distractor

Butir_Soal	isi_Distrac	Frekuensi	Proporsi	tus_Distrac
S1	A	4	0.111	Berfungsi
S1	C	2	0.056	Berfungsi
S1	D	4	0.111	Berfungsi
S1	E	6	0.167	Berfungsi
S2	A	3	0.083	Berfungsi
S2	B	2	0.056	Berfungsi
S2	D	2	0.056	Berfungsi
S2	E	2	0.056	Berfungsi
S3	A	4	0.111	Berfungsi

Hasil analisis menunjukkan bahwa sebagian besar distractor pada butir soal S1–S20 **berfungsi dengan baik**, ditandai dengan proporsi pemilihan $\geq 5\%$ oleh peserta. Hal ini menunjukkan bahwa opsi pengecoh mampu menarik peserta yang belum menguasai materi. Namun, terdapat beberapa distractor yang **tidak berfungsi**, yaitu pada butir S6, S13, S16, S17, dan S19 karena memiliki proporsi pemilihan sangat rendah atau tidak dipilih sama sekali. Distractor tersebut perlu direvisi agar lebih logis dan efektif, sementara secara umum kualitas distractor dalam tes ini tergolong **baik**.

Untuk hasil analisa di excel dapat di download melalui link github: https://raw.githubusercontent.com/arikaps/analisisoal/main/hasil_analisis_butir_lengkap.xlsx

KESIMPULAN

Berdasarkan hasil pengembangan dan analisis instrumen soal pilihan ganda, dapat disimpulkan bahwa instrumen yang dikembangkan telah memenuhi kriteria kualitas yang baik secara teoretis dan empiris. Hasil validasi ahli menggunakan koefisien Aiken's V menunjukkan seluruh butir soal berada pada kategori relevan hingga sangat relevan dengan nilai antara 0,75–1,00, sehingga instrumen dinyatakan memiliki validitas isi yang sangat baik.

Hasil analisis empiris terhadap jawaban siswa menunjukkan bahwa tingkat kesukaran butir soal umumnya berada pada kategori sedang. Analisis daya pembeda mengindikasikan bahwa sebagian besar butir soal mampu membedakan kemampuan peserta didik dengan baik, meskipun terdapat beberapa butir yang perlu direvisi atau dibuang karena daya pembeda rendah. Reliabilitas instrumen berdasarkan KR-20 sebesar 0,607 berada pada kategori sedang, yang menunjukkan konsistensi pengukuran yang cukup baik. Analisis distractor menunjukkan bahwa sebagian besar pengecoh telah berfungsi secara efektif, walaupun masih terdapat beberapa distractor yang perlu diperbaiki. Secara keseluruhan, instrumen soal dinyatakan valid,

cukup reliabel, dan layak digunakan, serta analisis menggunakan Python terbukti efektif dalam mendukung evaluasi butir soal secara objektif dan sistematis.

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142.
<https://doi.org/10.1177/0013164485451012>
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan* (Edisi revisi). Jakarta: Bumi Aksara.
- Azwar, S. (2019). *Reliabilitas dan validitas* (Edisi 4). Yogyakarta: Pustaka Pelajar.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160. <https://doi.org/10.1007/BF02288391>
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). Boston: Pearson Education.
- Sudijono, A. (2015). *Pengantar evaluasi pendidikan*. Jakarta: RajaGrafindo Persada.
- Sugiyono. (2019). *Metode penelitian pendidikan: Pendekatan kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- McKinney, W. (2018). *Python for data analysis: Data wrangling with pandas, NumPy, and IPython* (2nd ed.). Sebastopol, CA: O'Reilly Media.
- Aiken, L. R. (1985). *Three coefficients for analyzing the reliability and validity of ratings*. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan* (Edisi revisi). Jakarta: Bumi Aksara.
- Azwar, S. (2019). *Reliabilitas dan validitas* (Edisi 4). Yogyakarta: Pustaka Pelajar.